UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

TOWARDS THE IDENTIFICATION OF BIOMARKERS FOR CYSTIC FIBROSIS BY PROTEOMICS

NUNO MIGUEL ANTUNES GARCIA CHARRO

DOUTORAMENTO EM BIOLOGIA

ESPECIALIDADE BIOLOGIA MOLECULAR

2011

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

TOWARDS THE IDENTIFICATION OF BIOMARKERS FOR CYSTIC FIBROSIS BY PROTEOMICS

Tese orientada pela Doutora Deborah Penque e Professora Doutora Ana Maria Viegas Gonçalves Crespo

NUNO MIGUEL ANTUNES GARCIA CHARRO

DOUTORAMENTO EM BIOLOGIA

(BIOLOGIA MOLECULAR)

2011

The research described in this thesis was conducted at Laboratório de Proteómica, Departamento de Genética, Instituto Nacional de Saúde Dr. Ricardo Jorge (INSA, I.P.), Lisbon, Portugal; Clinical Proteomics Facility, University of Pittsburgh Medical Centre, Pennsylvania, USA; and Laboratory of Proteomics and Analytical Technologies, National Cancer Institute at Frederick, Maryland, USA.

Work partially supported by Fundação para a Ciência e a Tecnologia (FCT), Fundo Europeu para o Desenvolvimento (FEDER) (POCI/SAU-MMO/56163/2004), FCT/Poly-Annual Funding Program and FEDER/Saúde XXI Program (Portugal). Nuno Charro is a recipient of FCT doctoral fellowship (SFRH/BD/27906/2006).

Agradecimentos/Acknowledgements

"Nothing is hidden that will not be made known;
Nothing is secret that will not come to light"

Desde muito pequeno, a minha vontade em querer saber mais e porquê foi sempre presença constante. Ao iniciar e no decorrer da minha (ainda) curta na investigação científica, as perguntas foram mudando, o método também e várias pessoas contribuíram para o crescimento e desenvolvimento da minha personalidade científica e pessoal. Espero não me esquecer de ninguém e, se o fizer, não é intencional; apenas falibilidade.

A realização deste trabalho não seria de todo possível sem a grande contribuição da Doutora Deborah Penque. A sua orientação, empenho na minha formação científica e as possibilidades que me abriu no seu laboratório permitiram dar mais um contributo na problemática da Fibrose Quística. Obrigado pela recepção, acompanhamento, vontade de aprender e entusiasmo que demonstrou nas várias etapas do trabalho. A oportunidade que me deu em colaborar com outros excelentes laboratórios abrindo portas para que este trabalho fosse mais sólido e tivesse a perspectiva de outros que o olhassem "de fora" foi, sem dúvida, fundamental. Realço também o seu lado humano, não de "chefe chata" mas de orientadora preocupada e que sabe perceber quando as coisas correm menos bem.

À Prof.ª Ana Maria Crespo por todo apoio institucional prestado assim como a confiança e entusiasmo depositados neste trabalho e confirmados pelos sucessivos convites para a apresentação no âmbito do Mestrado que coordena na FCUL.

Um obrigado especial à equipa de Pneumologia do Hospital de Santa Maria em Lisboa, coordenada pelo Prof. António Bugalho de Almeida, em especial à Dr.ª Pilar Azevedo e Dr. Carlos Lopes pelo interesse em participar na investigação científica, no recrutamento e caracterização dos doentes de Fibrose Quística e pela disponibilidade demonstrada para deslocações ao INSA quando este trabalho estava ainda no início. A investigação científica e a prática clínica têm tudo a ganhar com colaborações sólidas, interessadas e empenhadas nas suas diversas vertentes.

A todos os doentes e suas famílias por nos aceitarem na privacidade da consulta de rotina, pela partilha de informações e disponibilidade para, mais uma vez, participarem na recolha de amostras biológicas nem sempre de forma tão indolor quanto se apregoa. Espero que este trabalho venha um dia a beneficiar-vos a todos pois, em última instância, ele é para vós. As melhoras e tenham toda a esperança possível num futuro melhor!

Ao grupo do Prof. Francisco Couto da FCUL e aos seus alunos, em especial à Cátia Pesquita e ao Daniel Faria, pelo tempo passado em frente ao computador a ajudar-nos a responder a perguntas que às vezes nem sequer sabíamos formular. As vossas "ferramentas" permitiram tornar a análise de resultados mais fácil e rápida. Espero que esta tese contribua para as divulgar a tantas pessoas quanto possível e aproveitar o seu potencial.

Obrigado ao INSA nos seus vários departamentos mas em especial a todos no Departamento de Genética por me terem recebido sempre tão bem, pela simpatia diária nos corredores e pelos bons momentos que passámos nas nossas festas de Natal.

Aos meus colegas e amigos de laboratório: Patrícia, Bruno e João. Pelas nossas (meias) torradas e meias de leite matinais quando a vontade era pouca ou quando o entusiasmo e a pressa nos faziam engoli-las à pressa. A companhia e amizade não são traduzíveis em palavras e vários são os momentos passados que me deixam a rebentar de orgulho de ser vosso amigo. Relembro em especial o "Prada Gang" quando a noite já ia bem longa nas nossas maratonas de colheita de amostras ou o "Green Day" quando, mesmo sem palavras, aparecemos todos vestidos da mesma cor. Obrigado por crescerem comigo, obrigado por me deixarem crescer convosco.

À Isabel que foi a minha primeira mentora de bancada naqueles dias mesmo antes do Natal em que o desafio de começar um projecto novo com pessoas que ainda não conhecia se misturava com o medo de poder não corresponder às expectativas.

Às mais recentes "aquisições" do laboratório: Fátima, Tânia, Sofia e Vukie. Apesar de "metidas todas no mesmo saco", cada uma de vocês tem aquela contribuição única para

o nosso grupo. Fui aprendendo a ouvir alguma da vossa música sem ter palpitações ou ansiedade no peito.

Aos meus amigos, todos aqueles que me acompanharam nesta demanda e que, como eu, também se debatiam (debatem?) sobre o papel da investigação e qual será afinal o nosso lugar nela. As aulas na FCUL afastaram-nos mas sabe sempre bem quando nos reunimos todos para voltar a recordar o passado e olharmos para o futuro.

À minha família no geral e aos meus pais e irmã no particular. Obrigado por nunca me terem fechado quaisquer portas e caminhos por explorar, por manterem em aberto todas as minhas opções. Obrigado por terem aturado os meus momentos mais entusiasmantes quando talvez vos apetecesse perguntar "Mas de que raio está ele a falar?!" ou quando não me apetecia falar em nada e mesmo assim divergíamos a conversa para outro assunto de forma a deixar-me voar.

Numa fase mais final, ao Filipe por me ter permitido encontrar um equilíbrio, um balanço na vida e dar-me aquela esperança de que amanhã será melhor e que amanhã seremos melhores.

Obrigado a todos os que passaram e continuam a passar pela minha vida nas suas várias vertentes. É da conjugação de todas elas que sou hoje alguém melhor e mais atento. No final de contas, sou E.U.

List of Publications

Sousa PM, Silva ST, Hood BL, Charro N, Carita JN, Vaz F, Penque D, Conrads TP, Melo AM, *Supramolecular organizations in the aerobic respiratory chain of Escherichia coli*, Biochimie. 2010 Oct 30, PMID: 21040753

Charro N, Hood BL, Faria D, Pacheco P, Azevedo P, Lopes C, de Almeida AB, Couto FM, Conrads TP, Penque D, *Serum proteomics signature of cystic fibrosis patients: a complementary 2-DE and LC-MS/MS approach*, J Proteomics. 2011 Jan 1;74(1):110-26, PMID: 20950718

Hood BL, Grahovac J, Flint MS, Sun M, Charro N, Becker D, Wells A, Conrads TP, *Proteomic analysis of laser microdissected melanoma cells from skin organ cultures*, J Proteome Res. 2010 Jul 2;9(7):3656-63, PMID: 20459140

Carvalho-Oliveira IM, Charro N, Aarbiou J, Buijs-Offerman RM, Wilke M, Schettgen T, Kraus T, Titulaer MK, Burgers P, Luider TM, Penque D, Scholte BJ, *Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model*, J Proteome Res. 2009 Jul;8(7):3606-16, PMID: 19438287

Santos MR, Rodríguez-Gómez MJ, Justino GC, Charro N, Florencio MH, Mira L, *Influence of the metabolic profile on the in vivo antioxidant activity of quercetin under a low dosage oral regimen in rats*, Br J Pharmacol. 2008 Apr;153(8):1750-61, PMID: 18311191

Simões T, Charro N, Faria D, Chan KC, Isaaq HJ, Couto FM, Waybright T, Veenstra TD, Blonder J, Penque D, *Proteomic Analysis of Human Nasal Epithelium: a Molecular Portrait*, Submitted to Journal of Proteomics Special Issue, 4th EuPA Meeting 2010 Proteomics Odyssey Towards Next Decades

Charro N, Simões T, Faria D, Chan KC, Isaaq HJ, Couto FM, Waybright T, Veenstra TD, Blonder J, Penque D, *Deep Proteome Profiling of Nasal Epithelial Cells:*

Consequences for Impaired Respiratory Function in Cystic Fibrosis Patients, Manuscript under final preparation

Charro N, Hood BL, Fonseca A, Zerimech F, Hagenfeldt M, Gomes F, Martins F, Neto D, Conrads TP, Peneda J, Penque D, *Proteome profiling of immunodepleted serum from individuals undergoing an alcoholic abstinence program*, Submitted to Journal of Proteomics Special Issue, 4th EuPA Meeting 2010 Proteomics Odyssey Towards Next Decades

List of Abbreviations

α-CHCA: α-Cyano-4-hydroxycinnamic acid
2-DE: Two-Dimensional Electrophoresis
ΔF508: Deletion of the phenylalanine (F) residue in position 508
Ab: Antibody
ABC family: ATP-Binding Cassette transporters family
ACN: Acetonitrile
ASL: Airway Surface Liquid
ATP: Adenosine Triphosphate
BALF: Bronchoalveolar Lavage Fluid
BMI: Body Mass Index
bp: base pair
CaCC: Calcium activated Chloride Channel
cAMP: Cyclic Adenoside Monophospate
CBAVD: Congenital Bilateral Absence of Vas Deferens
CBB: Coomassie Brilliant Blue
cDNA: Complementary Deoxyribonucleic Acid
CF: Cystic Fibrosis
CFRD: Cystic Fibrosis Related Diabetes
CFTR: Cystic Fibrosis Transmembrane Conductance Regulator
CID: Collision-induced Dissociation
CIP: CFTR Interacting Proteins
Cl^-: Chloride ion
COPD: Chronic Obstructive Pulmonary Disease
DIGE: Difference in Gel Electrophoresis
emPAI: Exponentially Modified Protein Abundance Index
ENaC: Epithelial Sodium Channel
ER: Endoplasmic Reticulum
F508del: Deletion of the phenylalanine (F) residue in position 508
FA: Formic Acid
FDR: False Discovery Rate
FEV_1: Forced Expiratory Volume in 1 second
FVC: Forced Volume Capacity
GO: Gene Ontology
GSH: Reduced Glutathione
GSSG: Oxidized Glutathione
GTP: Guanosine triphosphate
HCO_3^-: Bicarbonate ion
HSP: Heat-shock Proteins
IEF: Isoelectric Focusing
IL: Interleukin

IPG: Immobiline pH Gradient
kb: kilobase
kDa: kilo Dalton
LC: Liquid Chromatography
LIT: Linear Ion Trap
LPS: Lipopolysacharide
MALDI-TOF: Matrix Assisted Laser Desorption/Ionization-Time of Flight
MARS: Multiple Affinity Removal System
MBL: Mannose-binding Lectin
mEq/L: milliequivalents *per* liter
MHC: Major Histocompatibility Complex
mNEC: membrane fraction of Nasal Epithelial Cells
mRBC: Red Blood Cells' membrane
mRNA: Messenger Ribonucleic Acid
MRM: Multiple Reaction Monitoring
MS: Mass Spectrometry
MS/MS: Tandem Mass Spectrometry
MSD: Membrane Spanning Domain
MUC: Mucin
MW: Molecular Weight
Na^+: Sodium ion
NaCl: Sodium Chloride
NEC: Nasal Epithelial Cells
NBD: Nucleotide Binding Domain
NFkB: Nuclear Factor kappa B
NO: Nitric Oxide
ORCC: Outwardly Rectifying Chloride Channel
PAGE: Polyacrylamide Gel Electrophoresis
PAI: Protein Abundance Index
PBS: Phosphate-buffered solution
PCL: Pericilliary Clearance
PK: Protein Kinase
ppm: parts *per* million
PTC: Premature Termination Codons
PTM: Post-translation Modifications
R: Regulatory Domain
RNS: Reactive Nitrogen Species
ROS: Reactive Oxygen Species
SARS: Severe Acute Respiratory Syndrome
SDS: Sodium Dodecyl Sulfate
SELDI-TOF: Surface Enhanced Laser Desorption/Ionization- Time of Flight
sNEC: soluble fraction of Nasal Epithelial Cells
SRM: Single Reaction Monitoring
TPC: Total Peptides Count

TFA: Trifluoroacetic Acid
TGF: Transforming Growth Factor
TNF: Tumor Necrosis Factor
TMD: Transmembrane Domains
TOF: Time of Flight
TS: Total Signal
UPC: Unique Peptides Count
UPR: Unfolded Protein Response
WB: Western Blot

List of Figures

List of Supplemental Data

Table II.1SD: Statistical comparison given by Progenesis SameSpots of the 52 differentially expressed protein spots identified, having as reference their expression in Controls (**II.1.1SD**) and of the 37 differentially expressed protein spots identified, having as reference their expression in CF patients with mild lung disease (**II.1.2SD**).

Table II.2SD: MS data of the 52 and 37 protein spots analysed (Swiss-Prot DB 2008.06.09, 385721 sequences; 19630 human sequences) according to the mutational-based (**II.2.1SD**) and respiratory-based analysis (**II.2.2SD**).

Table II.3SD: Post-translation modifications predicted by FindMod on the identified proteins by 2DE.

Table II.4SD: Total proteins identified by label-free 1D-RPLC-MS/MS and their respective peptide count for each group under analysis.

Table II.5SD: Peptides' sequences obtained by-1D-RPLC-MS/MS that contributed to the identification of proteins in the immunodepleted fraction of serum.

Table II.6SD: Significant molecular functions, biological processes and cellular compartment retrieved by Gene Ontology (GO) of the differentially expressed proteins in immunodepleted serum.

Table II.7SD: Trends of expression of selected proteins identified in immunodepleted serum.

Figure II.1SD: Functional characterization of differentially expressed proteins in the immunodepleted fraction of serum according to Ingenuity Pathway Analysis.

Table III.1SD: PIKE annotation of mRBC proteins identified.

Table III.2SD: Significant molecular functions, biological processes and cellular compartment retrieved by Gene Ontology (GO) of the differentially expressed proteins in mRBC.

Table IV.1SD: Total proteins, their correspondent peptide count and the respective peptides' sequences identified by label-free 2D-LC-MS/MS in soluble and membrane fractions of NEC (sNEC and mNEC, respectively).

Table IV.2SD: Proteome annotation on sNEC and mNEC according to different databases.

Table IV.3SD: Hydrophobicity analysis of the sNEC and mNEC subproteomes (TMD and GRAVY).

Table IV.4SD: Functional characterization of sNEC and mNEC subproteomes according to Ingenuity Pathway Analysis.

Table IV.5SD: Estimation of total protein content and contribution of individual protein's amount in a complex mixture by emPAI.

Table IV.6SD: Comparison of NEC and bronchial proteomes.

Table IV.7SD: Association between the identified proteins in NEC and respiratory diseases.

Table V.1SD: Data normalization by the Total Signal approach (TS) and PIKE annotation on proteins used for differential studies in CF.

Table V.2SD: Data normalization by the emPAI approach and PIKE annotation on proteins used for differential studies in CF.

Table V.3SD: Significant molecular functions, biological processes and cellular compartment retrieved by Gene Ontology (GO) of the differentially expressed proteins in mNEC.

Table V.4SD: Functional analysis of the differentially expressed proteins in mNEC according to IPA.

Table V.5SD: Significant molecular functions, biological processes and cellular compartment retrieved by Gene Ontology (GO) of the differentially expressed proteins in sNEC.

Table V.6SD: Functional analysis of the differentially expressed proteins in sNEC according to IPA.

Figure V.1SD and V.2SD: Functional analysis of the differentially expressed proteins in mNEC according to IPA.

Figure V.3SD and V.4SD: Functional analysis of the differentially expressed proteins in sNEC according to IPA.

Table of Contents

THEMATIC RANGE OF THE WORK AND KEYWORDS

The scope of the work presented in this thesis is to provide identification and characterization of proteins that might contribute to the development and/or progression of the pulmonary disease Cystic Fibrosis (CF), aiming to a better understanding on the proteinaceous environment in several biological samples with influence in the disease. The objectives of this work were in accordance with the actual and to-date Portuguese National Health Plan and were developed in the National Institute of Health Dr. Ricardo Jorge in Lisbon, Portugal, in collaboration with the University of Pittsburgh Medical Centre and the National Cancer Institute at Frederick, both in the United States of America. The involvement of such institutions recalls the importance of this study in trying to unravel mechanisms underlying the development and/or progression of this treatable but incurable disease.

This thesis is divided in six chapters which contain results already published or under final preparation for submission to peer reviewed scientific journals of the area, according to point 1 of Artigo 41 from Capítulo V of Regulamento de Estudos Pós-Graduados da Universidade de Lisboa, published in Diário da República, 2ª Série - Nº 209 de 30 de Outubro de 2006. In this context, I hereby state that all the experiments, results and interpretation contained here are original and performed by myself with the collaboration of others where mentioned.

The **Chapter I** corresponds to the **General Introduction**, where several aspects of the CF's pathology are approached, since its recognition as an independent disease to the several CF-causing mutations and the phenotype evidenced by the patients. The contribution of several others molecular determinants other than the CFTR alone and the role of Proteomics in these achievements is also described.

Experimental design, methodologies, results and interpretation are provided in chapters II to V, each one dedicated to a different type of biological sample.

Chapter II devotes it to the proteomics characterization of human serum by complementary approaches and integration of the identified species into networks and pathways with biological significance in the context of CF.

Chapter III describes the characterization of the CF-associated proteome of mRBC and aims to interpreter previous phenomena reported in these cells of CF patients by the study of altered abundances levels of some relevant proteins.

Chapter IV provides a wide molecular portrait of proteins being expressed in the nasal epithelial cells ultimately demonstrating the usefulness of this easily collectable biological sample in mimicking lungs' microenvironment and its application to the study of respiratory diseases.

Chapter V makes use of the deep proteome profiling of nasal epithelial cells and establish correlations and consequences for the impaired respiratory function observed in CF patients.

Finally, a **General Discussion** is presented in **Chapter VI** were the main conclusions and remarks are pointed hoping to unveil new perspectives on CF while open novel questions to be answered by Proteomics.

Keywords: Cystic Fibrosis; Proteomics; Pulmonary Function; Bioinformatics; Functional Characterization; Pathway Analysis.

RESUMO DA TESE

A Fibrose Quística (FQ) é a doença monogénica autossómica recessiva mais frequente e letal na população caucasiana, com uma incidência de 1 em cada 2500-3200 recém-nascidos sendo que 1 em cada 25-30 indivíduos apresentam pelo menos uma mutação no gene responsável pela patologia. Esta doença é causada por mutações no gene *Cystic Fibrosis Transmembrane Conductance Regulator* (*CFTR*) que codifica para uma proteína com o mesmo nome e cuja principal função é servir de canal de cloreto regulado pelo cAMP (do Inglês, *cyclic Adenosine Monophosphate*) e cujos ciclos de abertura e fecho dependem do ATP (do Inglês, *Adenosine Triphosphate*) na membrana apical das células epiteliais. Actualmente, conhecem-se mais de 1800 mutações no gene da *CFTR* mas a delecção de um resíduo de fenilalanina na posição 508 da cadeia polipéptídica (F508del; ΔF508) constitui cerca de 70% dos cromossomas FQ analisados e está presente em cerca de 90% dos doentes de FQ, originando um produto com problemas ao nível da maturação e, consequentemente, disfuncional. Defeitos na síntese, estrutura, processamento e/ou função da CFTR ditam uma multiplicidade de sintomas observados nestes doentes: elevadas concentrações de cloreto e sódio no suor, insuficiência pancreática, disfunção intestinal e hepática, infertilidade e obstrução das vias respiratórias por um muco espesso e desidratado, criando as condições propícias ao desenvolvimento de infecções respiratórias crónicas e um estado de inflamação constante. De facto, a maior parte dos doentes FQ vê a sua qualidade de vida significativamente diminuída pela deterioração progressiva da função pulmonar que, em última instância, conduz à morte. Vários estudos têm sido efectuados no sentido de estabelecer uma relação entre o genótipo apresentado pelos doentes FQ e as manifestações clínicas da doença, tarefa dificultada pela heterogeneidade de mutações e sintomas apresentados pelos doentes bem como pela influência que o ambiente e outros factores celulares (denominados de genes modificadores da FQ) exercem. Dessa forma, importa conhecer que entidades celulares são estas cuja expressão e função poderão influenciar o desenvolvimento e/ou progressão da FQ, podendo mesmo constituir alvos terapêuticos da doença. Uns dos candidatos óbvios a este lugar são as proteínas. O seu estudo através das várias metodologias disponibilizadas pela Proteómica tem permitido um grande avanço na caracterização e conhecimento dos mecanismos envolvidos no processamento da CFTR; importa agora conhecer e investigar todo o ambiente proteico que contribui para o desenvolvimento e progressão da patologia.

O trabalho apresentado nesta tese surge como uma continuação da investigação que tem vindo a ser desenvolvida no Laboratório de Proteómica do Departamento de Genética do Instituto Nacional de Saúde Dr. Ricardo Jorge que se iniciou com o estudo dos mecanismos envolvidos no processamento e maturação da CFTR em linhas celulares, seguindo depois para um modelo animal tentando identificar proteínas diferencialmente expressas em pulmão e sua relevância para o desenvolvimento da doença pulmonar até ao estudo de amostras biológicas humanas no sentido de melhor caracterizar e compreender outras proteínas que não apenas a CFTR cujas alterações nos seus níveis de expressão poderão contribuir para a elucidação dos mecanismos da patologia. Neste trabalho, foram analisados soro, membranas de eritrócitos e células do epitélio nasal de doentes de FQ em comparação com indivíduos saudáveis portadores ou não de uma mutação no gene *CFTR*, através de várias metodologias tecnológicas da Proteómica.

Os objectivos foram atingidos com recurso a tecnologias existentes no Laboratório de Proteómica em Lisboa, como por exemplo a separação de proteínas por 2-DE (do Inglês, *Two-Dimensional Electrophoresis*) e sua identificação por MS (do Inglês, *Mass Spectrometry*), mais propriamente por MALDI-TOF/TOF (do Inglês, *Matrix-Assisted Laser Desorption/Ionization-tandem Time of Flight*) em colaboração com outros laboratórios nos EUA (*Clinical Proteomics Facility, University of Pittsburgh Medical Centre, Pennsylvania e Laboratory of Proteomics and Analytical Technologies, National Cancer Institute at Frederick, Maryland*) que nos permitiram fazer uso de tecnologias de ponta tanto na separação robusta como na identificação confiante de proteínas por MudPIT (do Inglês, *Multidimension Protein Identification Technology*).

No capítulo II, são apresentadas e discutidas as funções de proteínas diferencialmente expressas que poderão ter implicações para a doença no soro dos vários grupos de indivíduos estudados. A razão deste estudo prende-se com o elevado potencial do soro para a identificação de biomarcadores por conter um elevado número de proteínas provenientes de todos os órgãos do organismo, ser de fácil colheita e constituir uma amostra biológica por excelência para a implementação de testes clínicos. A inovação deste estudo centra-se na utilização de técnicas complementares de separação e identificação de proteínas, nomeadamente 2-DE-MALDI-TOF/TOF MS e

MudPIT. Por 2-DE foram identificadas 28 proteínas únicas a partir de 78 spots de gel enquanto que por LC-MS/MS o número de proteínas identificadas aumentou para 569. A utilização de diversas ferramentas bioinformáticas permitiu ainda alocar estas proteínas a processos com significância biológica e cuja regulação tem sido amplamente descrita em FQ: remodelação dos tecidos, desequilíbrio na razão proteases/antiproteases, inflamação, estado nutricional e stresse oxidativo ou disfunção imune, entre outros. A estratégia aplicada demonstrou ser eficiente na elucidação de processos e no conhecimento mais alargado de proteínas activamente implicadas na patogénese da FQ que é corroborado por diversos estudos já publicados.

Para além de estudar a componente líquida do sangue, pretendeu-se também estudar parte da sua componente sólida, as membranas de eritrócitos, uma vez que estudos têm sugerido o papel destas células no desenvolvimento de hipertensão pulmonar em FQ, embora de uma forma pouco clara. A presença da CFTR nos eritrócitos tem sido demonstrada por alguns autores através da modulação da sua actividade e/ou por métodos imunológicos. Os nossos dados confirmam a presença da CFTR na membrana dos eritrócitos e constituem, portanto, o primeiro estudo que reporta a sua identificação por MS nestas células. Contudo, esta identificação só foi possível em eritrócitos de doentes homozigóticos para a ΔF, provavelmente por haver maior acumulação desta proteína *misfolded* pelo controlo de qualidade da célula. O estudo da caracterização do proteoma presente nas membranas dos eritrócitos de doentes FQ revelou um número de proteínas diferencialmente expressas associadas a processos que poderão estar na base do desenvolvimento da hipertensão pulmonar nestes doentes: desregulação e desequilíbrio na manutenção da forma e deformabilidade dos eritrócitos, com consequências ao nível da incapacidade de libertação de ATP e produção do vasodilatador pulmonar óxido nítrico (NO), e aumento do seu stresse oxidativo com implicações no aumento da susceptibilidade para a peroxidação lipídica das membranas e redução das defesas contra agressões externas, podendo levar, em última instância, à deterioração da função pulmonar nestes doentes. O conjunto de resultados aqui obtidos permite oferecer um conhecimento mais consubstanciado sobre processos já elucidados mas cujos cujas proteínas associadas estavam ainda caracterizadas.

Sendo o pulmão o principal órgão afectado na FQ, torna-se claro que a amostra biológica desejável para a condução de estudos para descoberta de biomarcadores será uma biopsia deste tecido. No entanto, e por diversas razões entre as quais se destacam as éticas e as clínicas, o acesso a este tipo de amostras é muito restrito em doentes e/ou mesmo impossível quando se trata de indivíduos controlos. Várias amostras como lavado nasal e brônquico ou expectoração têm sido usadas com este objectivo, não representando contudo as características celulares do tecido uma vez que constituem fluidos segregados. A utilização das células do epitélio nasal obtidas por escovagem da cavidade nasal tem sido considerada um bom modelo biológico por mimetizar as condições das vias respiratórias inferiores e cuja colheita é deveras mais fácil. Dado que o proteoma do epitélio nasal humano estava ainda pouco caracterizado, no Capitulo IV é apresentado um estudo proteómico exaustivo das proteínas que o constituem. Através da aplicação de metodologias de pré-fraccionamento celular para extracção de proteínas solúveis e membranares, combinadas com análise de larga escala através de MudPIT, foram identificadas mais de 7000 proteínas das quais 1482 foram consideradas com confiança como características e representativas do epitélio nasal. Comparações de resultados com diversas bases de dados permitiram a identificação de várias funções associadas à manutenção e função de um epitélio estruturado, na percepção e resposta a agressões externas, assim como uma significativa sobreposição entre proteínas identificadas no epitélio brônquico e neste epitélio nasal e sua correlação com diversas doenças respiratórias, entre as quais cancro do pulmão, síndroma respiratório agudo grave, pneumonite, asma e mesmo FQ. A conjugação das várias observações apoia a hipótese que o epitélio nasal reflecte certos aspectos das vias respiratórias inferiores e contribui significativamente para o incremento de conhecimento do fenótipo do epitélio respiratório, permitindo análises comparativas em condições patológicas como a FQ.

A informação resultante da caracterização do proteoma das células do epitélio nasal serviu como base para identificação de proteínas diferencialmente expressas como consequência de FQ, apresentada no Capítulo V. A identificação de processos relacionados com a remodelação dos tecidos, o stresse oxidativo e a inflamação foram identificados neste epitélio como associados a FQ. Também foram evidenciados a disfunção mitocondrial e desregulação nos mecanismos de produção energética, metabolismo da glucose ou degradação da matriz extracelular, entre outros. A

identificação de proteínas diferencialmente expressas e que interactuam com a CFTR e a identificação/participação de proteínas conhecidas como modificadoras da FQ reveste-se também de fulcral importância para patogénese pulmonar da FQ.

Em suma, os resultados obtidos nas várias amostras biológicas de doentes FQ estudas, por diferentes abordagens proteómicas, permitiram obter conhecimentos únicos que consideramos de base sólida para uma melhor compreensão da patologia FQ e possibilitar o desenho de estudos posteriores e dedicados a proteínas/funções elucidadas com o objectivo de validação de alguns potenciais biomarcadores de diagnóstico/prognóstico/monitorização e/ou alvos terapêuticos específicos que permitam distinguir a FQ de outras doenças respiratórias.

Chapter I

General Introduction

1. A Timeline for Cystic Fibrosis

The history of Cystic Fibrosis (CF) could start with an "A long, long time ago…" since the first records of this disease go back to the 17th century in German and Swiss folklore recognizing the association between salty skin and early death: "Woe to that child who tastes salty from a kiss on the brow, for he is cursed and soon must die" [1]. Also, several descriptions of infants and children with *meconium ileus* and typical CF's pancreatic and pulmonary disease are reported since 1650. In the 19th century, Karl Landsteiner first described *meconium ileus* and Carl von Rokitansky associated it with foetal death in a CF inborn [2]. It would take another century till CF was recognized as a distinct clinical entity by Guido Fanconi and Dorothy Hansine Anderson by first describing the characteristic "cystic fibrosis of the pancreas" characterized by mucous plugging of the glandular ducts of pancreas from children dying of malnutrition and correlating it with lung and intestinal disease representative of CF [3,4]. Moreover, she also hypothesized that CF was a recessive inherited disease and started the usage of pancreatic enzyme replacement to ameliorate the dysfunction of exocrine pancreas.

Another designation for CF is mucoviscidosis, a term introduced by Farber in 1945 and still in use in non-English-speaking countries to describe a "generalized state of thickened mucus" as consequence of CF [5]. In 1953, Paul Di Sant'Agnese discovered abnormalities in sweat electrolytes in CF patients during an unusual heat wave in New York, reporting elevated levels of Cl^- and Na^+ [6], fact that was on the basis of the development of an accurate and affordable diagnostic method that is still the most common CF diagnostic tool [7]. This was further demonstrated by Knowles and colleagues who described imbalanced transport of Cl^- and Na^+ in CF respiratory epithelium [8], and Quinton *et al.*, that uncovered impermeability of sweat gland ducts to Cl^- in CF [9], postulating that a transporting channel was malfunctioning in CF's epithelia.

It was only in 1989 that a team led by Lap-Chee Tsui identified the gene responsible for CF and its transcript, detected in sweat glands, lung and pancreas [10-12]. This remarkable discovery opened the possibility to uncover the molecular basis of the disease while giving the opportunity to explore other molecular determinants that might contribute to the development and/or progression of CF.

2. Cystic Fibrosis: Clinical Description

CF is the most common life-limiting autosomal recessive disease in Caucasians. Although the disease frequency varies among different ethnic groups, it affects about 1 in every 3200 newborns and has a carrier frequency of 1 in 25 individuals [13,14]. This monogenic disease is caused by mutations in the Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*) gene that codes for the CFTR protein, a cAMP-dependent and ATP-gated Cl^- channel. More than 1800 mutations have been described to be CF-causing though the deletion on phenylalanine at position 508 (F508del; ΔF) is the most frequent, being present in around 70% of CF chromosomes and in 90% of heterozygotes [15] (http://www.genet.sickkids.on.ca/cftr). Defects in CFTR synthesis, structure and/or function as a Cl^- channel in the apical membrane of several epithelia account for a multiplicity of clinical complications that characterize CF: high sweat Cl^- and Na^+ concentrations, excessive dehydrated mucus secretion that causes exocrine pancreatic insufficiency, neonatal *meconium ileus*, male infertility due to congenital bilateral absence of the vas deferens (CBAVD) and obstruction of the airways. In fact, impaired mucocilliary clearance in the lungs resulting from mucus obstruction leads to recurrent respiratory infections by *Pseudomonas aeruginosa* and *Staphylococcus aureus*, bronchiectasis and ultimately death [1,16,17]. Some of the symptoms can be overcome by pancreatic enzymes replacement or nutritional supervision. However, persistent respiratory infections originate an exacerbated inflammatory response in the airway epithelia resulting in loss of pulmonary function and are still the main cause of morbidity and mortality in CF even under intensive antibiotic therapies.

Earlier diagnosis, such as prenatal and newborn screening [18,19], and comprehensive and integrated clinical treatments allowed an increase in the mean life expectancy of CF patients from 8 years in 1974 to almost 50 in present days [15,20,21] though it is still a life-threatening disease [22]. The standard diagnosis of CF is the sweat test in which $[Cl^-]$ in sweat must be > 60mEq/L on at least two repeated measurements [23,24] and genotyping for mutations in the *CFTR* gene. However, not all patients present the typical CF-associated complications and even diagnosis can be dubious with borderline sweat Cl^- levels, indicating that factors other than mutated CFTR may be involved and contribute to the phenotype.

3. From the *CFTR* gene to CFTR protein

A new chapter in CF research started in 1989 with the identification of the *CFTR* gene. Chromossome walking and jumping allowed the identification of the gene in the long arm of chromosome 7 band 31 (7q31.2), spanning approximately 250 kb and comprising 24 coding exons [12]. Its transcript is 6.5 kb and was found to be largely restricted to epithelial tissues, though in low levels [11,25]. Sequence comparison between alleles from patients and their parents revealed the major CF mutation, a deletion of 3 bp in exon 10 causing a loss of phenylalanine at position 508 of the protein: ΔF or, more recently, F508del [10] (**Figure I.1**).

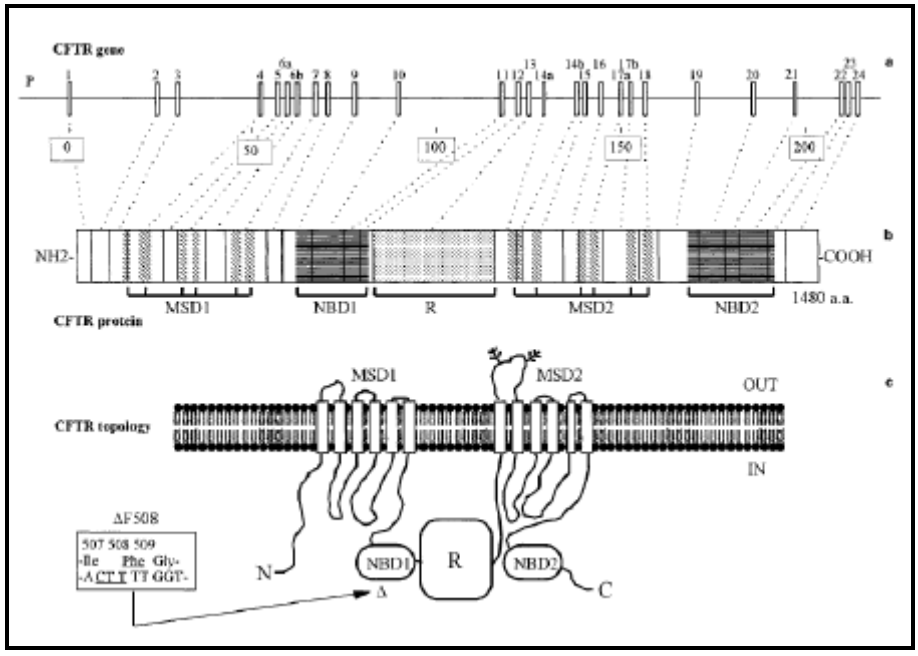


Figure I.1: From the *CFTR* gene to CFTR protein. On top, structure of the *CFTR* gene with the promoter region (P) followed by the 27 coding exons; in the middle, linear structure of the 1480 amino acids that constitute the polypeptidic chain of CFTR protein showing predicted domains (MSD – membrane spanning domain; NBD – nucleotide binding domain; R – regulatory domain); in bottom, CFTR topology and its distribution through the membrane, intra- and extracelular compartments. Deletion of three nucleotides with consequent loss of phenylalanine at position 508 (F508del, the most frequent mutation in CF's chromosomes at NBD1) is highlighted (box). (Adapted from [26]).

3.1 CFTR protein

The product of the *CFTR* gene is a single polypeptide chain of 1480 aminoacids with apparent molecular mass of 140 to 170 kDa, depending of its glycosylation status and type of cell where it is expressed [11,27]. CFTR's splicing variants have been

reported though no functional significance has been reported so far [28,29]. The protein is composed of two highly hydrophobic membrane spanning domains (MSDs), each consisting of six α-helices (transmembrane domains, TM1 to TM12) and five loops, which form the channel pore; two nucleotide binding domains (NBDs) capable of ATP hydrolysis and one large regulatory domain (R) with multiple phosphorylation sites that are phosphorylated primarily by cAMP-dependent protein kinase A (PKA) but also by protein kinase C (PKC) [27,30]. According to this topology, the vast majority of the protein is cytoplasmic (both NBDs and R domain lie in the cytoplasm), part is integrated in the membrane and only a small portion is extracelular [31-33]. However, it was only in 2008 that CFTR's 3D structure was obtained and resolved, answering to many questions related with CFTR's topology and function, thus opening frontiers to new questions and studies [34,35] (**Figure I.2**).

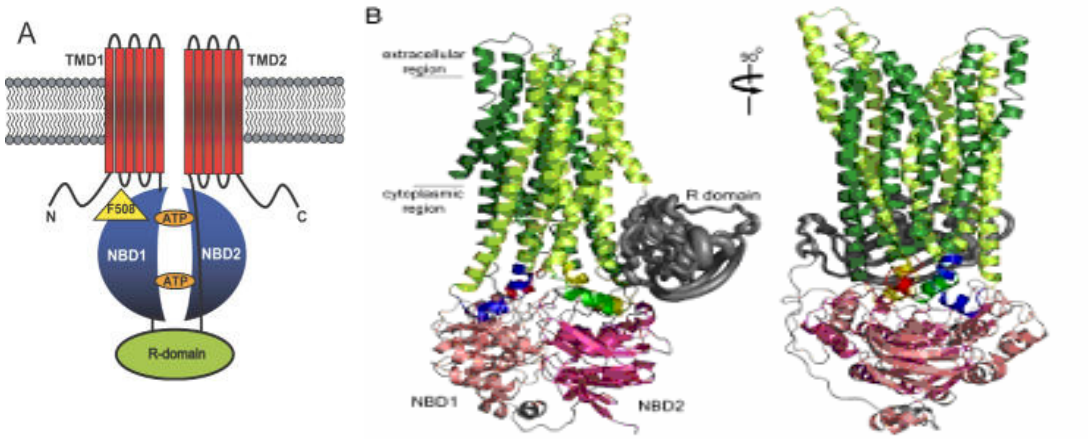


Figure I.2: **(A)** Schematic view of the domain structure CFTR protein. TMD – transmembrane domain; NBD – nucleotide binding domain; R – regulatory domain [34]. **(B)** The 3D structure of CFTRconstructed by molecular modelling and supported biochemically. (Adapted from [34,35]).

CFTR belongs to the ATP-binding cassette (ABC) superfamily of transporters, a class of proteins found in both prokaryotes and eukaryotes responsible for the transport of a wide variety of substrates across biological membranes at the ATP hydrolysis expense [36] but that lack the R domain. In the CFTR case, ATP hydrolysis is needed for conformational alteration of the protein and regulated channel gating [30,37].

3.2 Localization

Sites of expression and signals that modulate CFTR expression can provide critical insights into the CF pathogenesis, CFTR functions as well as to develop appropriate and specific gene therapy or to therapeutically manipulate endogenous CFTR expression. The initial description of *CFTR* gene was also accompanied by transcript analysis highlighting important sites of expression [11].

The predominant sites of CFTR expression are many of the epithelia surfaces throughout the body, most of them corresponding to CF-affected tissues: submucosal glands, airway epithelium [38], pancreatic ductal epithelium [39], epithelium along the gastrointestinal tract [39], sweat glands [40], epithelium of the developing genital ducts, adult epididymis, vas deferens [41,42], cervix and uterus [39,42], salivary glands [39,43] and the epithelium of the intrahepatic bile ducts and gallbladder [44]. However, some epithelial sites of CFTR expression do not evidence the disease such as kidney collecting duct epithelium [45] or the submucosal glands of duodenum [43]. Besides being expressed in epithelial cells, several reports claim CFTR expression in several non-epithelial tissues such as neuronal expression in brain [46], aortic smooth muscle [47,48], ventricular cardiomyocytes [49,50], lymphocytes [51] and erythrocytes [52,53], among others, although in significantly lower values. The relevance of CFTR expression in such tissues is unclear.

In addition to have a distinctive pattern and cell-specific expression, CFTR is also regulated in a development-dependent manner. The most well-known site of developmental regulation of CFTR expression is the airway surface epithelium, with relatively high expression during embryonic and foetal development with significant decrease at birth, due to airway epithelium switch from secretory to absorptive [41,54]. In fact, CFTR mRNA and protein levels are markedly less abundant in lung and airways than in other epithelia. CFTR predominant site of expression in the human large airways is the submucosal glands [38] but mRNA and protein levels have been detected in distal airway, bronchioles and alveoli [55] and nasal epithelia [56-58].

Intracellularly, CFTR will traffic through the components of the secretory pathway (endoplasmic reticulum, Golgi apparatus, endosomal and lysossomal parts) during migration to apical plasma membrane [59-61].

3.3 Functions

3.3.1 CFTR as a Chloride Channel

Once the structure and conformation of a protein is established, inferences about its function can be assumed. Even before the identification of the *CFTR* gene, it was already known that CF resulted from an abnormally low permeability to Cl^- in sweat gland ducts resulting in a high concentration of NaCl in sweat [9]. The first experimental approach to understand CFTR's function was the transfection of CF epithelial cells with CFTR wild-type cDNA to restore the cAMP-mediated Cl^- conductance [62-64]. The most convincing proof that CFTR is a Cl^- channel was performed by Bear *et al.* after incorporation of purified CFTR into artificial lipid bilayers that exhibited regulated Cl^- channel activity [65]. CFTR is an ohmic low-conductance channel (conductance $\leq$10 pS) and, accordingly to the model currently accepted, the mechanism of gating is tightly regulated by ATP levels and cellular balance of kinase and phosphatase activity. PKA and PKC phosphorylate serine residues within the R domain upon cAMP-dependent activation. When this happens, channel gating is regulated by a cycle of ATP hydrolysis at the NBDs. Dephosphorylation of the R domain leads the channel returns to its "inactive" state [30]. Phosphorylation of R domain alters its conformation and interaction with other parts of the protein. The 3D structure of CFTR (Figure I.2) reveals important aspects of the interfaces between the NBDs and MSDs of CFTR, which are crucial in both assembly during biogenesis and in mediating conformational signals that influence channel activity [34]. In spite of significant advances in the understanding of CFTR structure and function, some aspects of CFTR regulation are still controversial and not fully understood [34,66,67].

In brief, the main functions of CFTR accounts for endogenous cAMP-regulated and ATP-gated Cl- channel activity mostly at the apical membrane of epithelial cells.

3.3.2 CFTR as Regulator of Other Channels

Transport of Cl^- is in close correlation to the transport of its counter ion Na^+. Although fluid and electrolyte secretion is decreased in CF due to CFTR transport dysfunction, hyperabsorption of Na^+ and hyperactivity of Na^+ channels is observed across the CF airway epithelia [68]. Stutts *et al.* described that activation of CFTR downregulates epithelial Na^+ channel (ENaC) and that is missing in CF [69,70]. The

results are accumulation of mucus, pulmonary congestion, reduced mucociliary clearance and excessive dryness of the epithelial surface.

In human airways, the activity of Ca^{2+}-activated Cl^- conductance channels (CaCC) has been described as being inhibited in the presence of CFTR and upregulated in CF. These channels mediate an increase in intracellular Ca^{2+} and interact with the C-terminal of R domain of CFTR [71]. Though the CaCC inhibition by CFTR does not require CFTR activation, it is potentiated when CFTR is activated.

Another example of Cl^- channels regulated by CFTR are the outwardly rectifying Cl^- channels (ORCCs), some of the most abundant channels in epithelial cells and defective in CF. Activated CFTR facilitates the release of extracelular ATP which in turn stimulates Cl^- conductance of ORCCs [72,73].

Renal outer medullary potassium (K^+) channel (ROMK2) and delayed activated voltage dependent K^+ channel (KvLQT-1), K^+ channels located at basolateral membrane, are activated by CFTR [74,75]. Extracellular conductance of K^+ is essential to ensure the driving force for Cl^- secretion.

Studies in Xenopus oocytes have shown that overexpression of CFTR increases water permeability through aquaporin-3 (AQP3) [76]. This protein was also identified in respiratory epithelial cells with consequent increase in water permeability under CFTR activation [77,78].

3.3.3 Other Functions of CFTR

Besides being a regulator of several ionic channels, other functions have been attributed to CFTR. Moreover, defects and/or dysfunction of CFTR solely as a Cl^- channel do not completely explain the complex phenotype observed in CF patients.

CFTR is required for the transport of bicarbonate (HCO_3^-) in the lungs, pancreas and gastrointestinal tract. The implication of bicarbonate in CF pathophysiology can be explained by its ability to act as a pH buffer and enhancing the solubility of proteins and ions in biological fluids. Ion selectivity studies show that CFTR is 3 to 5 times more selective for Cl^- than to HCO_3^- [79] and that it transports HCO_3^- in a process mediated by coupled Cl^--dependent transport [80]. Bicarbonate is highly relevant in CF especially at the pancreatic level: CFTR mutants with normal or substantial Cl^- channel activity but which do not transport HCO_3^- are associated with more severe forms of pancreatic

disease, while those that show reduced HCO_3^- transport are associated with milder forms [81].

Glutathione (GSH) is a tripeptide that constitutes the most important extracelular antioxidant in the lungs and is implicated in the regulation of inflammation and immune response [82]. CFTR was shown to directly mediate the nucleotide-activated transport of GSH. In CF, the GSH concentration in airway surface liquid (ASL) that covers the respiratory epithelium is greatly reduced [83], providing evidence that CFTR can play a role in GSH flux from epithelial cells to ASL and, consequently, in the control of oxidative stress and chronic inflammation in the airways [84].

It is known that airway cells in CF present an altered glycophenotype that dramatically favours the chronic colonization by *P.aeruginosa* and *Haemophilus influenzae* [85]. Binding to the epithelial cells activates the host immune response resulting in an exacerbated inflammatory state of the airways mucosa and deterioration of pulmonary tissue in CF. In addition to that, these two pathogens that most commonly colonize the CF lungs have binding proteins that recognize the altered glycoconjugates in CF cells [86]. CFTR tends to accumulate in the specific region of the membrane where bacteria bound and can be assumed that the protein is needed for signalling in response to infection [87] and even in the removal of the infector [88]. The glycosylation pattern is also important in compartmentalization and transport of the terminal glycosyltransferases through the Golgi. Mutated/absent CFTR causes incorrect compartmentalization which leads to modified glycophenotypes in the surface of CF cells' membranes and compromised bacteria clearance [89].

CFTR also seems to regulate exocytosis especially in polarized epithelial cells [85], process that is defective in CF. Its internalization from the apical membrane is inhibited by an increase in cAMP levels and dependent of CFTR itself as a Cl^- channel [90]. Endosome fusion also seems to be CFTR- and cAMP-dependent [91].

4. Classification of CFTR Mutations

As previously mentioned, over 1850 sequence variations (mutations and polymorphisms) have been reported in the *CFTR* gene (http://www.genet.sickkids.on.ca/cftr) stressing the need of classification and grouping according to their effect on CFTR function. Zielinski an Tsui [92] have proposed the classification of CFTR's mutation in five categories (**Figure I.3**):

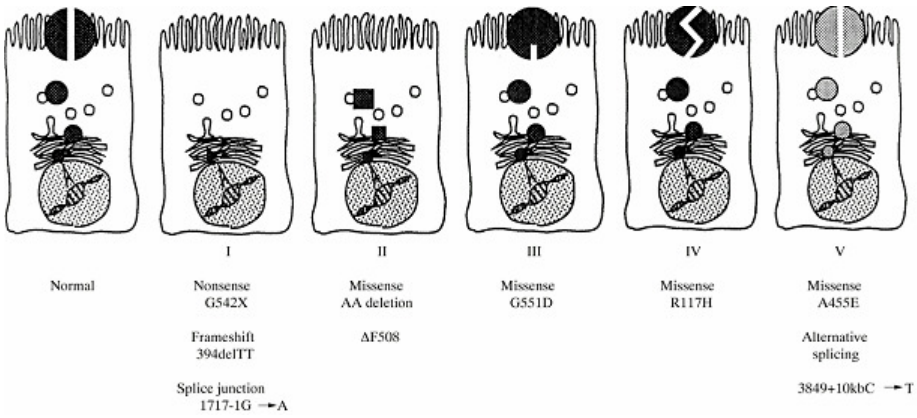


Figure I.3: Classification of CFTR mutation and its molecular consequences. (Adapted from [92]).

● **Class I - Defective Protein Synthesis:** Mutations in this class correspond to almost half of the studied mutations and include nonsense, splicing and frameshift mutations that lead to creation of premature termination codons (PTC) resulting in complete absence of CFTR protein. PTC generally result in truncated proteins; however, additional effects on mRNA instability have been reported. The final consequence is a more severe form of CF [93]; *e.g.*, G542X, W1282X, R553X.

● **Class II – Defective Protein Processing:** Mutants carrying class II mutations present a defective processing and/or intracellular trafficking in CFTR maturation with absent or reduced protein concentration at the plasma membrane. The disease severity correlates with the amounts of fully processed CFTR that is able to reach its site of action; *e.g.*, F508del, N1303K.

● **Class III – Defective Protein Regulation:** In this class, normal levels of CFTR protein are produced and inserted into the plasma membrane but are unable to function as a cAMP-activated Cl^- channel due to inability to be activated by ATP or cAMP. Phosphorylation/dephosphorylation is the major pathway by which CFTR is regulated and it is defective in this class of mutations. Patients carrying such mutations present a wide range of phenotypes from severe lack of CFTR function to reduced response to cAMP stimulation and slightly reduce activity; *e.g.*, G551D, Y569D, S1255P.

● **Class IV – Altered Conductance:** Mutations of class IV are located in regions corresponding to the TMDs that form the channel pore affecting the gating properties and causing a reduced Cl^- efflux rate. Here again, normal levels of CFTR

protein are produced, processed, transported and inserted into the apical membrane but display a defective conductance though with residual activity; *e.g.*, R347P, R334W.

● **Class V – Reduced CFTR Level:** This class includes promoter mutations that reduce transcription and amino acid substitutions that cause inefficient protein maturation and, consequently, reduces the amount of protein in the membrane; *e.g.*, 3272-26A>G, G576A, 3849+10kb C→T.

Vankeerberghen *et al.* [94] have proposed a sixth category corresponding to alterations in CFTR that result in protein forms with reduced stability and increased degradation rate with consequences in the functioning of other channels and imbalanced electrolyte movement across membrane [95].

Mutations in classes IV, V and VI frequently cause milder forms of the disease (patients are usually pancreatic sufficient) as partially functional CFTR is present at the membrane [96].

Despite the enormous number of mutations and polymorphisms, one mutation accounts for approximately 70% of the CF chromosomes and 90% of at least one allele in CF patients: F508del. This mutation results in a shortened-one-amino-acid protein in position 508 of the polypeptide chain with altered misfolding, incomplete glycosylation and incorrect trafficking to plasma membrane [97-99]. The other four most common mutations are G542X, G551D, N1303K and W1282X, with the remaining being located in geographically restricted areas.

The genetic and physiological information of the *CFTR* gene promoted a better understanding of the pathology; however, the levels of functional protein are the central point in determining the manifestations of CF: the lower the levels of active CFTR, more organs become affected with more severe consequences. The opposite is also verified **(Figure I.4).**

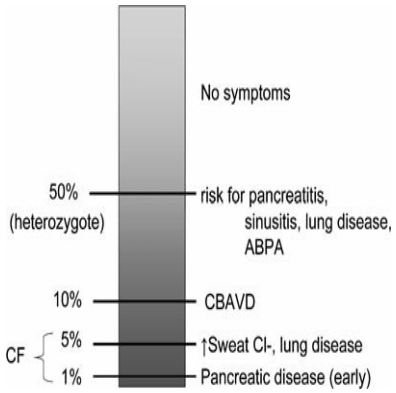


Figure I.4: Correlation between the level of active CFTR and the phenotype observed in patients (Adapted from [21]).

5. CF Phenotype

Classic CF is characterized by progressive obstructive lung disease, pancreatic insufficiency and congenital absence of the vas deferens in males while non-classic forms of the disease account for the non-appearance of the classic triad of symptoms [19]. In any of the cases, multiple organs are affected with consequences in the welfare and life quality of the CF patients (**Figure I.5**).

5.1 Sweat Glands

Sweat testing by quantitative pilocarpine iontophoresis is the main standard test for confirming the diagnosis of CF [19,100,101]. Sweat is produced at the base of the sweat gland as a salty fluid and flows to the skin surface through a narrow duct. In the final part of the duct, ions are absorbed making the sweat water-rich. Epithelial tissues in CF patients are unable to absorb Cl^-, and consequently Na^+, from the duct lumen causing sweat to retain an excess of Cl^- and Na^+, becoming abnormally salty [102]. Quantitation of Cl^- concentrations is required, though [Na^+], conductivity and osmolarity may also be measured. The advent of the development of such sweat test was in 1949 during a heat wave in New York City due to enormous increase in hospitalization of CF patients that rapidly became dehydrated [6]. Today, it is established that for the sweat test to serve as diagnostic of CF, it must have [Cl^-] > 60mmol/L in at least two measurements, performed preferably several weeks apart, and be interpreted in an age-specific fashion [19,100,103].

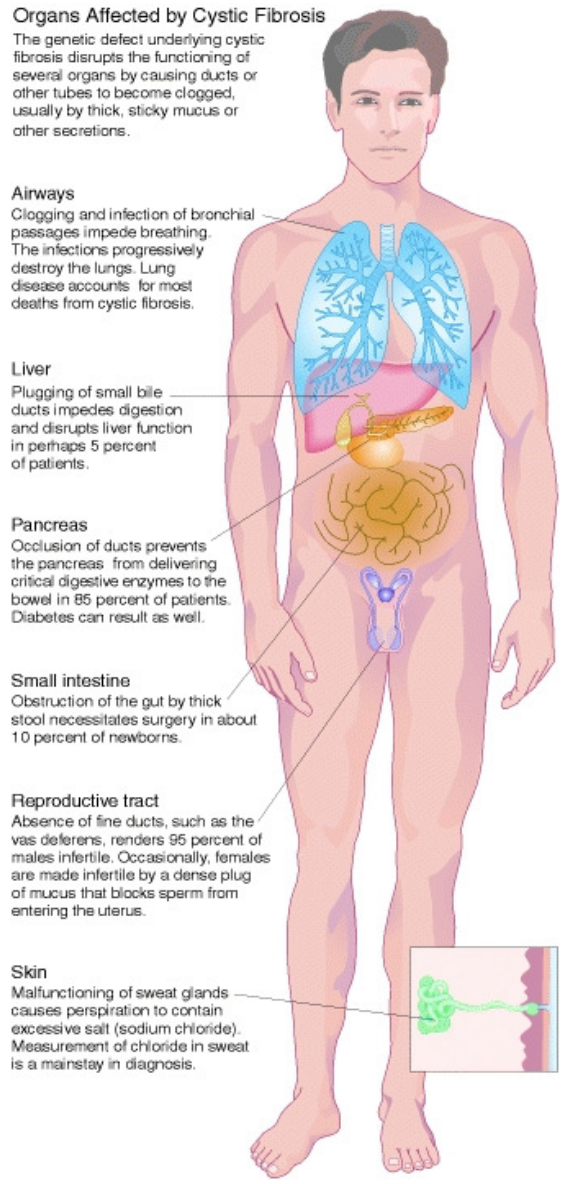


Figure I.5: Consequences of dysfunctional CFTR in human organs. (Adapted from http://www.ncbi.nlm.nih.gov).

5.2 Intestinal tract

Gastrointestinal dysfunction remains the earliest and most common manifestation of CF with distal intestinal obstruction syndrome, *meconium ileus* and rectal prolapse being prematurely observed even in prenatal sonography [104]. There are some similarities between the fluid and electrolyte secretion in the intestinal tract and in the airways: Cl^- enters the basolateral membrane via co-transport processes and exits apically via specific channels such as CFTR. In CF, the Cl^- secretory component is absent [105] leading to alterations in the composition of the luminal content. The final

consequence is nutrients' malabsorption in the gut, malnutrition, reduced body mass index (BMI) in patients and increased propensity for recurrent infections.

5.3 Hepatobiliary disorders

CFTR is expressed in epithelial cells of the biliary tract, namely in cholangiocytes and gallbladder epithelia. About one third of CF patients show abnormal results of liver function presenting features such as hepatomegaly, diffuse fatty infiltration to severe cirrhosis. Histological changes include intraluminal concretions in the biliary tree with duct dilatation. Inflammation and fibrosis also occur due to bile-duct epithelium hyperplasia and proliferation [106], similar to that observed in the pancreas. In addition, intrahepatic biliary epithelial cells produce excessive mucus composed of proteoglycans, which increase the viscosity of CF bile [107]. A small, poorly functioning gallbladder affects up to 30% of CF patients and gallstones up to 10% [108].

5.4 Fertility

98% of CF men are infertile, with aspermia, secondary to atretic or congenital bilateral absence of vas deferens (CBAVD) and dilated or absent seminal vesicles [109]. The vas deferentia become blocked with viscid secretions early in life and are reabsorbed [110] but sexual potency and spermatogenesis are kept in a normal level [20]. Females with CF are also found to be less fertile than normal healthy women. This is mainly caused by tenacious impermeable cervical mucus, which does not undergo the typical changes during menstrual cycle, due to defective CFTR protein expressed in the cervix [111,112]. Congenital absence of the uterus and vagina (CAUV) are also commonly associated with CF [113]. The incidence of CFTR mutations found in the patients with CAUV was twice as high as that found in the general population (8% vs 4%), but much less than the incidence of CFTR mutations in men with CBAVD (80%). This data suggests that it is unlikely for CFTR mutations to cause CAUV in females as they cause CBAVD in males [113].

5.5 Pancreatic Disease and CF-related Diabetes (CFRD)

Exocrine pancreatic insufficiency is present in about 90% of CF patients and constitutes the most-well studied and established genotype-phenotype correlation (see 6.

Genotype-Phenotype Correlation). Patients with mild *CFTR* mutations that have almost normal levels and activity of exocrine pancreatic enzymes are called pancreatic sufficient while those who present severe *CFTR* mutations do not have such enzymes and need to supplement their meals with enzymatic supplements are called pancreatic insufficient [114]. Pancreatic disease results from retention and prematurely activation of digestive proenzymes in the pancreatic ducts due to a reduction in volume of pancreatic secretion fluid lower in HCO_3^- concentration, leading to tissue destruction and fibrosis [20,115] (**Figure I.6**). The resulting malabsorption fails to meet the raised energy demands caused by hypermetabolic state with endobronchial infection causing a vicious cycle of reduced appetite, vomiting, malnutrition and infection. Diabetes mellitus are rare during the first decade of the patients' life as Langerhans cells are initially spared from destruction and fibrosis. Incidence of CFRD increases with age [108,116]. Currently, pancreatic insufficiency can be corrected by enzyme replacement [117].

5.6 Respiratory Disease

Chronic lung disease is the major cause of morbidity and mortality in CF and the link between ion transport defects and the pathogenesis of CF lung disease has been the subject of major clinical research. Along with pancreatic disorder and sweat glands dysfunction, pulmonary disease forms the core of classic CF.

It is generally accepted that innate host defence is deeply altered in the disease that include chemical protection by antimicrobial peptides/proteins such as lysozyme or lactoferrin and mechanical clearance of mucus and inhaled particles and/or pathogens by mucociliary apparatus [118,119] (**Figure I.6**).

So far, several hypotheses have been proposed aiming to explain the mechanisms between defective CFTR and airway disease: High Salt hypothesis [120]; Low pH hypothesis [121]; Low Oxygenation hypothesis [122,123]; Low Airway Surface Liquid (ASL) Volume hypothesis [124] and Abnormal Gland hypothesis [125]. These assumptions are based on the facts that increased Na^+ absorption and reduced ability of the CF airways to secrete Cl^- alters ASL volume, promote mucus adherence and favour bacterial colonization and infection. Adding to all of this, it is predicted that CF pathogenesis reflects deficient epithelial binding and phagocytosis of bacteria [88,126] or an intrinsically enhanced inflammatory response in CF airways [127]. Persistent and recurrent bacterial infections trigger an inflammatory response creating a

vicious cycle of infection, inflammation, impaired mucociliary clearance and oxidant/antioxidant status that promotes bronchiectasis, respiratory failure and ultimately death [128].

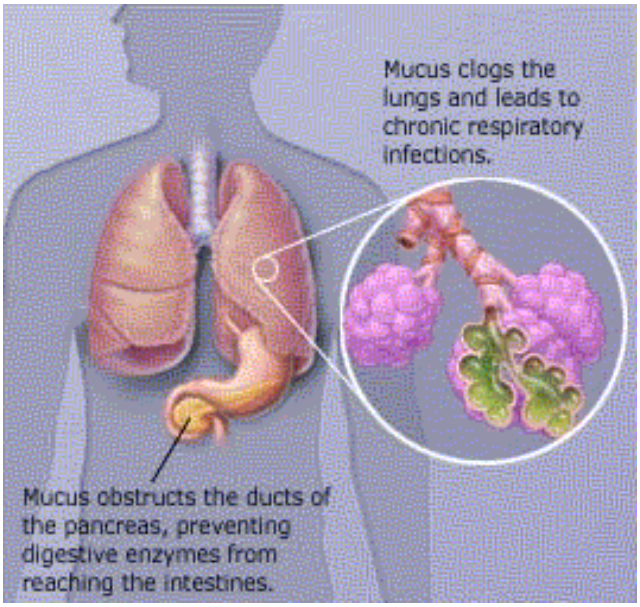


Figure I.6: Pathogenesis of CF at pancreas and lung level. (Adapted from http://www.medicine.ufl.edu).

5.6.1 Airway Surface Liquid (ASL) and Periciliary Liquid Layer (PCL)

CF patients are born with apparent no lung disease, though mucus plugging, abnormal ASL volume and infection are responsible for degrading lung function early in life [127]. The airway epithelia is highly water permeable and the ASL volume is determined by the mass of salt on airway surfaces, which in turn is tightly regulated via two opposing active transport systems, ENaC-mediated Na^+ absorption and Cl^- secretion mediated by CFTR and CaCC [129,130]. In CF, increased ENaC activity and defective CFTR function were predicted to deplete ASL volume. ASL can be divided into two compartments: a mucus layer, where inhaled particles and pathogens are entrapped, and the periciliary liquid layer (PCL), a mucin-free low viscosity medium that facilitates ciliary beating and that functions as lubricant to separate mucus layer from the mucins tethered to the cell surface to facilitate cough clearance [119,129]. In contrast to CF epithelium, normal airway epithelia have the capacity to regulate the volume of ASL by setting the height of the PCL to the height of an extended cilium. Moreover, PCL height in CF is significantly reduced and cilia are flattened/collapsed onto cell surfaces [129,131-133]. This alters optimal mucociliary transport and creates

an unfavourable microenvironment for antibacterial subjects, such as lysozyme and defensins, to perform their activity [134].

All these events provide a significant link between abnormal ion transport and deficient mucus clearance in the airways of CF patients.

5.6.2 Pathogens

Shortly after birth, the lungs of CF patients become infected by a variety of pathogens, namely *Staphlylococcus aureus* and *Haemophilus influenza* [135,136]. *Pseudomonas aeruginosa* affects approximately 80% of the population [15] and is associated with a more rapid clinical deterioration, especially mucoid strains [135,137]. Other species infecting CF patients' lungs account for *Stenotrophomonas maltophilia* and *Burkholderia cepacia* [15]. (**Figure I.7**) Although great development and improvement in antibiotic treatment strategies against respiratory-tract infections, these are regarded as the main reason of morbidity and mortality [138,139].

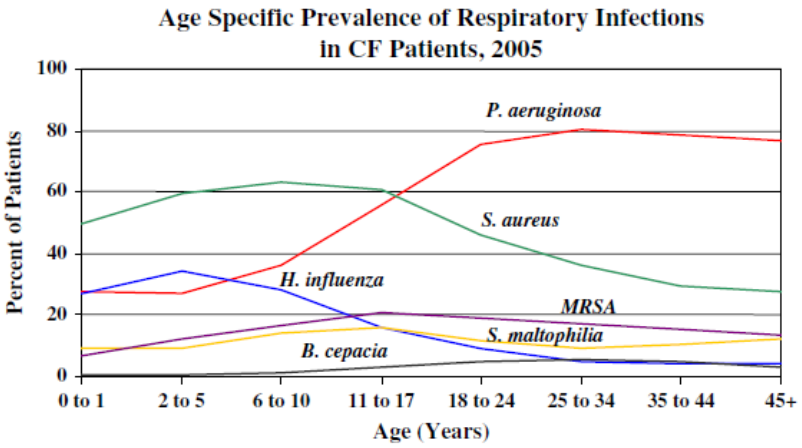


Figure I.7: Specific pathogens tend to colonize CF's patients' lungs in an age-dependent manner (Adapted from [21] using data from the 2005 Cystic Fibrosis Patient Registry, Bethesda, MD, USA).

Currently, several hypotheses aim to explain the development and persistent colonization of CF's airways by opportunistic pathogens: Inflammation-first hypothesis [140]; Cell-receptor hypothesis [141]; Salt-defensin hypothesis [142,143] or the Isotonic fluid depletion/anoxic mucus hypothesis [131], though none of them completely elucidate about all the types of colonization and infection.

Pseudomonas aeruginosa is one of the hallmarks of CF and causes chronic deterioration and damage in lungs. Its environmental ubiquity and direct contact from patient-to-patient explain its high frequency [144]. Recent research has highlighted mechanisms of attachment of the microorganism and emphasizes the role of hair-like proteins (Pili) and flagella in the attachment, adhesion and biofilm formation in the CF lung. After an initial and transient colonization period with non-mucoid strains, untreated patients generally became chronically infected with alginate-coated mucoid strains [145] which became antibiotic-resistant due to low penetration of the drugs and rapid development of mutant strains [146].

Staphylococcus aureus (MRSA) are an important human pathogen, resistant to penicillin and frequently isolated from sputum samples of adult CF patients [147]. Although they do not significantly contribute to the morbidity and mortality in CF [148], it is suggested that MRSA-induced lung damage can predispose to *Pseudomonas aeruginosa* infection [149]. No significant change was observed in MRSA-infected patients regarding respiratory function, nutritional status or chest x-ray appearances [150] though a small increase in dyspnoea, wheeze and sputum production arises.

Burkholderia cepacia was first identified by Francis Burkholder in 1950 [151] and the first report in CF's airways is from late 1970's [152-154]. *Burkholderia cepacia* affects <10% of CF patients but its management is important in both clinical and surgical evaluation of patients selected for lung transplantation [155,156] as they are commonly resistant to antibiotics and host defensins [157].

Lungs of CF patients are constantly under attack of opportunistic microorganisms. Prevention of bacterial lung infection and effective antibiotic therapies are the primary aim for CF treatment with gains in quality and increased life expectancy.

5.6.3 Infection *versus* Inflammation

Inflammation and infection are two major components of the CF lung disease though whether which one takes place at first is still object of great debate.

The inflammation-first hypothesis proposes that inflammation is present in the airways of patients in the first steps of life before infection [140], findings supported by work in germ-free raised mice with CF [158]. Some reports on CF foetuses and newborns that show that their lungs appear histologically and morphologically normal although with distended mucus glands in the upper airways [15,159]. Others have observed increased

inflammatory cells and proinflammatory cytokines interleukin-8 (IL-8) and interleukin-6 (IL-6) in bronchoalveolar fluid (BALF) in infants with CF in the absence of apparent infection [135,140] without excluding the possibility of an anterior infection cleared at the time of testing. This *à priori* inflammatory phenotype is supported by at least other two groups that have reported increased production of proinflammatory cytokines and exaggerated NFkB activation in lung epithelial cells expressing mutant forms of CFTR [160,161].

On the other hand, evidences from Armstrong and co-workers do not agree with this hypothesis and emphasize that inflammation follows infection [136]. CF's BALF contain low concentrations of anti-inflammatory interleukin-10 (IL-10) produced from bronchial epithelial cells [162] that can predispose patients to severe lung inflammation after colonization with *Pseudomonas aeruginosa* [163]. IL-10 modulates the inflammatory functions of monocytes/macrophages, lymphocytes and neutrophils in the airways [164] and inhibits transcription factor NFkB, a pathway that is highly activated in CF [165-167]. Other activators of NFkB accounts are reactive oxygen species (ROS) that are produced in the CF airways as consequence of pathogens' colonization [168]. Pier and coworkers proposed that CFTR is a pattern recognition molecule that extracts lipopolysaccharides (LPS) from outer membranes of *Pseudomonas aeruginosa* into the epithelial cells, activating NFkB signalling [169], and that ΔF-CFTR cells lack its initial proinflammatory response by IL-1-β-NFkB, resulting in chronic airway inflammation [170]. More recently, other mechanism of NFkB activation in CF was demonstrated to be the result of inability of CFTR to communicate with inflammatory receptors via TNFα protein complex gap junctions [170-172]. This phenomenon is mutation-type-dependent as ΔF508 promotes the activation of NFkB signalling 3.5 times more intense than G551D-CFTR [166].

In both "schools of thought", the NFkB proinflammatory signalling plays a central role in the development and progression of CF lung disease (recently reviewed in [173]) whether by increased predisposition for an inflammatory phenotype or by response to a bacterial colonization.

CF mouse models do not spontaneously acquire lung infection or display typical CF lung disease so their contribution in helping to clarify this matter is also limited. Investigators have studied xenografts of CF human tracheal and bronchial epithelial cells implanted in immunocompromised mice [174-176] though the results were

unclear, reinforcing the idea that current model systems are insufficient to prove whether inflammation can occur prior to infection.

Overall, an exaggerated, sustained and persistent neutrophilic inflammatory response to bacterial and viral pathogens, that ultimately destroy the airway's wall, is an accepted feature of lung disease in CF. It is assumed that, for a given bacterial load, a person with CF will have up to 10 times more inflammation than someone with lower respiratory tract infection but without the disease [177]. Inflammation is present even in young infants diagnosed by neonatal screening and in clinically stable patients [136,140,178].

6. Genotype-Phenotype Correlations

The classification of CFTR mutations on the basis of their molecular mechanisms represents an attempt to predict the course of the disease and to establish to what extent CF phenotype results from underlying mutations [114]. This task is however difficult as CF affects several organs and the symptoms are highly heterogeneous among different patients. Three factors seem to play an intricate role in defining CF phenotype: the specific pair of *CFTR* mutations, secondary genetic factors other than CFTR also called modifier genes and the environment [179].

There are two typical approaches to study genotype-phenotype correlations: to evaluate trends in patients carrying specific *CFTR* mutations and genotype analysis of similar clinical profiles, *e.g.*, pancreatic or lung disease, age of diagnosis, etc. Dominant phenotypic effects have been reported for mild *CFTR* alleles [180].

Zielenski and colleagues have predicted a good correlation between the class of a certain mutation and some clinical symptoms, though the severity of lung disease does not seem to be evident. On the other hand, the pancreatic function seems to correlate well with the *CFTR* genotype [26,92]. In general, patients with pancreatic insufficiency are homozygous or heterozygous of two mutations from class I, II or III and diagnosed later while those with pancreatic sufficiency have at least one mutated allele from milder classes IV or V and show symptoms in the earlier years of life. Malnutrition was significantly higher in patients with a severe genotype than in those carrying milder mutations [181]. A feature that appears to be almost genotype-independent is CBAVD that affects almost all male patients [182].

Lung is by far the most compromised organ in CF and therefore the main focus of investigation. Some degree of correlation between *CFTR* class of mutations and lung

disease have been proposed [93,183-190] where milder mutations account for better pulmonary function and lower rate of *Pseudomonas aeruginosa* colonization resulting in better survival than severe ones. Complex alleles, *i.e.*, alleles that have a second site mutation, were also associated with higher phenotypic variation [191-193].
CFTR mRNA transcript levels are being correlated with severity of selected functional and/or clinical parameters. A study from Ramalho *et al.* showed that as much as 5% of the normal level of wild-type CFTR mRNA could be sufficient to attenuate the development of severe lung disease [194].

Overall, the genotype-phenotype correlation studies have significantly contributed to our understanding of how and to what extent *CFTR* mutations contribute to the severity of the pathology although variability and context (organ)-dependent is observed. Also, the direct impact of the *CFTR* genotype, as disease trigger, seems to be limited to the initial phase of CF indicating that other factors than *CFTR* genotype alone play a role in the course of the disease.

7. CF Modifier Genes and CFTR Interactors

The complex and heterogeneous phenotype observed among CF patients cannot be completely explained by the impairment of CFTR alone, indicating that other factors might be implicated in this highly variable disease: the modifier genes [195]. Studies of twins and siblings affected with CF showed substantial impact of genetic modifiers in CF pulmonary disease, independent of CFTR genotype [196].

The *CFTR*-knockout mice provided the first concrete evidence of a CF-modifier gene. Although these mice do not evidence classical CF lung disease, they develop intestinal obstruction shortly after birth, usually leading to death, which resembles *meconium ileus* present in about 20% CF human patients [197]. Moreover, severity and lethality of intestinal obstruction in CF has been associated with variability in a CFTR-independent locus on mouse chromosome 7, though the specific gene has not been identified yet [198]. A similar gene locus has been detected on human chromosome 19 [199]. Identification of genes that modify CF pulmonary phenotype has been more challenging. Lack of animals that resemble CF lung disease required moving towards population-based studies investigating some logical genes, some of them that may improve or correct the underlying Cl^--transport abnormalities or modulate the cycle of infection and inflammation [200,201]. The first twin-based study tried to establish a correlation between lung function and body mass index (BMI), showing a higher

correlation in monozygous that in dyzigous twin pairs though independent analysis of these two features showed no significant differences [201]. Subsequent studies aiming to correlate lung function in CF with FEV_1, treatment and socioeconomic factors have also been evaluated (for review, see [202]). Clarke and colleagues proposed that an alternative Ca^{2+}-activated Cl^- channel would be present at the airways trying to overcome the lack of CFTR function [197]. Others include molecules involved in CFTR trafficking, alternative ion channels, inflammatory and anti-inflammatory mediators, antioxidants and mediators of airways reactivity (**Table 1**).

Table 1: Putative CF modifier genes. (Adapted from [203] and reviewed in [202]).

Category	Examples of Proteins
Inflammatory/Anti-inflammatory genes	Cytokines (TNF-α; IL-8; IL-10)
Ion/Water Transport	Alternative chloride channels; ENaC
Immunity/Airway Defense	MBL2; HLA locus; NOS
Proteases/Antiproteases	Neutrophilic elastase; α-1-antitrypsin; IFRD1
Post-translational processing/chaperones	Heat shock proteins (Hsp90; Hsc70)
Growth Factors	TGF-β
Oxidants/Antioxidants	Glutathione S-transferase
Airway Reactivity	β2-AR; EDNRA
Mucus/Mucins	Muc5AC

7.1 CFTR interactors

Proteins that interact with CFTR during its biosynthesis, processing and maturation play a role in regulating the CF phenotype, either by direct up- or down regulation of CFTR itself or by indirect processes, creating the correct environment for normal biogenesis and/ trafficking. Moreover, growing evidence has demonstrated that these interactions are spatial and time-dependent, affecting the processing, localization and the channel activity of CFTR within cells [204].

During co- and pos-translational folding, CFTR binds to several proteins such as Hsc70, Hsp 40, Hsp90, calnexin or ubiquitin ligating (E3) enzyme, among others [205-208] protecting the protein from aggregation, facilitating folding and aiding the

degradation of non-native conformers [209-211]. So far, Wang *et al.* provided the most complete-to-date CFTR interactors' list evidencing proteins involved in the molecular chaperone system, ER folding factors, enzymes from the peptidyl-prolyl cis-trans isomerases, proteasomal subunits, vesicle associated transport and cellular organization proteins [212].

Deeper knowledge on the CFTR interacting proteins will contribute to elucidate all the involved mechanisms from synthesis to membrane delivery and recycling, to study their influence in the CF's patophysiology and to modulate new therapeutic targets for CF therapy.

7.2 Growth Factors and Inflammatory genes as CF modifier genes

Polymorphisms in proinflammatory cytokines may play a crucial role in regulating the chronic lung inflammation in CF. Transforming growth factor beta (TGF-β) is a cytokine with both inflammatory and anti-inflammatory properties. In airway epithelium, TGF-β modulates fibroblast proliferation and collagen deposition [213-215] and polymorphisms in the TGFB1 gene lead to a reduced abundance of TGF-β, being associated with decreased risk of pulmonary fibrosis after lung transplantation [216,217]. In CF patients, it was observed that polymorphism in codon 10 of TGFB1 were associated with low production of the protein and slower deterioration of lung function [218,219] though unclear results were unravelled by Bartlett *et al.* [220]. The correlation of TGFB1 with asthma and COPD has also been demonstrated [221,222].

Tumor necrosis factor alpha, or TNF-α, is a proinflammatory cytokine found in high concentrations within the CF airway that might contribute to the neutrophil-mediated inflammatory response and has been inversely correlated with lung function [223,224]. A study from Hull e*t al.* showed that CF individuals carrying a polymorphism in the promoter region of the TNF-α gene had a significantly lower FEV_1 and poorer nutritional status [225]. Arkwright, however, found no association between this [226] and Low *et al.* found no link with any marker of severity using a larger cohort of patients [227]. Other variants of TNF-α have been identified but their role in CF is uncertain due to low reproducible studies [195].

Nitric oxide (NO) is synthesized by NO-synthetase 1 (NOS-1) in the lung that is often found with increased levels in the presence of inflammation. In CF, exhaled NO is however lower than expected and patients with NOS-1 variant alleles that lead to

particularly low levels of exhaled NO present increased risk for early colonization with *Pseudomonas aeruginosa* and *Aspergillus fumigatus* and non-specific inflammation of the airways. Increased number of GT repeats immediately upstream the transcription initiation site of the NOS-1 gene significantly correlated with the exhaled NO levels [228], predicting a slower decline in lung function during the 5 years follow-up period for CF patients who had the NOS-1 genotype associated with higher NO production in an age, genotype and bacterial colonization independent way. A new functionally important polymorphism (894G/T) affects the resistance of NOS-3 to proteolysis and that is induced by oestrogen in the vascular endothelium. The relation between NOS-1 and NOS-3 as independent modifiers or if there is a confounding effect of one upon the other remains unclear [229-231].

7.3 Antiproteases and Antioxidants as Modifiers of Respiratory Disease

Chronic inflammation of the airways results in an excess of destructive proteases such as neutrophil elastase, cathepsin G and proteinase 3 [232]. Although considered an acute phase protein, α-1-antitrypsin has an important function preventing this destruction acting as an antiprotease and anti-inflammatory agent in the lungs [233,234]. Patients with inherited deficiency in α-1-antitrypsin are at risk of emphysema and thus it was postulated that CF patients with co-existing α-1-antitrypsin deficiency would have a more severe pulmonary phenotype [235]. Recent results from Bartlett and co-workers associated the α-1-antitrypsin Z allele with greater risk of developing severe liver disease [220]. On the other hand, Mahadeva *et al.* have found that decreased levels of α-1-antitrypsin may have a milder CF phenotype [236].

As for antioxidant candidates to CF modifier genes, glutathione S-transferases (GST) are among the most well positioned. This family of antioxidant enzymes are present in high concentrations in the airway epithelial liquid layer and function in the detoxification of oxidative stress products [237]. GSTM1 gene has a common non-functioning allele (GSTM1-0) that does not produce a protein. This polymorphism is associated with decreased lung function growth in children [238], high risk for chronic bronchitis [239], greater disease severity accessed by chest radiograph and Shwachman score and decreased survival [225,240]. Polymorphisms in another member of this family, GSTP1, have been associated with susceptibility to asthma [241-243].

7.4 Antigen Presentation and Host Defense Mechanisms

The human leukocyte antigen (HLA) is the most polymorphic in the human genome, encoding hundreds of genes including the major histocompatibility complexes (MHC). MHC class II molecules are fundamental in antigen presentation and ensuing inflammatory response. Polymorphisms have been associated with autoimmune conditions, asthma and allergy. An allelic variant, DR7, has been associated with IgE levels and increased risk for *Pseudomonas aeruginosa* colonization in CF [244], while DR2 allele showed increase propensity for allergic bronchopulmonary aspergillosis in CF [245].

Mannose-binding lectin (MBL) is a member of the collectin family being an important component in the innate immune defense system and exerting its function both by direct opsonization of pathogens and by activation of the complement system [246]. MBL is a product of MBL2 gene with polymorphisms (*MBL2-0*) in the promoter region resulting in non-functional mutants and low levels in circulation [247]: heterozygoty for *MBL2-0* results in decreased serum MBL and homozygoty in undetectable levels of the protein. In CF, *MBL2-0* is a likely candidate as a modifier gene of the lung disease as both FEV_1 and FVC were significantly lower in *Pseudomonas aeruginosa* infected subjects with either one or two structural *MBL2* mutations and increased risk for *Burkholderia cepacia* infection [248,249]. They also had increased risk of lung transplantation and death. A study involving a larger cohort of individuals (almost 600 subjects) showed that adult patients with homozygous for two structural *MBL2* mutations had significantly impaired lung function, oxygen saturation and raised inflammatory markers [250]. Conflicting reports have been published regarding MLB deficiency in childhood: Muhlebach and co-workers reported that MBL deficiency may initially be associated with higher lung function [251] while more recently MLB2 deficiency was significantly associated with earlier infection with *Pseudomonas aeruginosa* and, consequently, with a more rapid decline of pulmonary function [252].

7.5 Airway Hyperresponsiveness

β2 adrenergic receptor (β2-AR) is a G-protein-coupled receptor present in the airways that stimulates adenylyl cyclase activity and is known to influence airway reactivity [253]. Polymorphisms in the β2-AR gene have been considered to influence severe asthma [254,255] and obesity [256]. Buscher and co-workers have conducted a

study with 126 CF patients where they studied the effect of three polymorphisms in the *β2-AR*, namely $Arg^{16}Gly^{16}$, and found a significantly lower FEV_1 in homo- or heterozygous for Gly^{16} when compared with homozygous for Arg^{16}. Also, Gly^{16} was significantly less frequent in the CF population, possibly implying a survival disadvantage. A rare polymorphism was identified by the team that showed a blunted cAMP response to isoproteronol stimulation and significantly decreased the pulmonary function [257].

Two other polymorphisms on codons 16 and 27 showed no association with lung disease [258-260].

Variants in the endothelin receptor type A (*EDNRA*) are also likely to contribute to CF lung pathology. Darrah and co-workers have observed an association between a variant in the 3'-UTR of *EDNRA* and ΔF-CFTR. *EDNRA* has been implicated in vasoconstrictive diseases altering smooth muscle function. Its involvement in CF lung disease might produce similar effects at the smooth muscle tone in the airways and/or vasculature [261].

Besides the ones presented here, several other candidate genes have been studied aiming to make a correlation between different CF phenotypes (for review, see [195], [262] and [202]). However, it seems highly unlikely that these genes by itself confer the typical CF phenotype. The effects of interactions between CF modifier genes with *CFTR* genotype, binary interactions other than CFTR and the environment will certainly promote more substantial effects and detectable features.

Among the most well-studied is the interaction between MLB2 either with CFTR [263] and with *Pseudomonas aeruginosa* [249]. Most of these CF modifier genes-CFTR interactions promote variant polymorphisms effects in the context of severe mutation homozygosity with rare exceptions [219]. Gene-gene interactions have been reported in CF as well as with pathogens and exposure to second-hand smoke (reviewed in [195]).

Several CF modifier genes have been identified by association through genome-wide studies significantly increasing the numbers of potential candidates. One of the challenges in this field is therefore the development of analytical tools to determine which interactions will have implications in the CF phenotype and its biological relevance. Proteomics, along with bioinformatics, is very well positioned in aiming to answer such pertinent questions.

8. Proteomics and its Significance to the Understanding of Cystic Fibrosis

Upon the discovery of the *CFTR* gene and its mutations, great expectations arose given the possibility to study and understand this complex monogenic disease. However, investigators soon realised that they were facing a much more intricate and multifaceted problem which investigation could benefit from combined technologies other than gene sequencing. The completion of the human genome project was a key point in the evolution of proteomics, the large-scale study of protein profiles under certain conditions and in a determined point in time [264,265]. Modern medicine nowadays focus its attention to the integrated examination of gene expression of the entire genome under physiological or pathological conditions and how its expression pattern defines and predicts the different phenotypes, progression and/or aggressiveness of a disease. Proteomics is increasingly pointed as a fundamental scientific discipline in which disease mechanisms are unravelled and novel biomarkers and therapeutic approaches are discovered in this post-genomic era [266].

Particularly, in CF-related research, proteomics appears as an almost logic complementary approach to investigate this disease where only one gene is affected/mutated but which consequences imbalance the functions of a multiplicity of organs. Although our understanding on the molecular determinants in CF have augmented significantly over the years, none of the approved treatments currently available are able to correct the CFTR defect and patients are still dying in consequence of the disease. To be able to define and compare protein profiles from CF patients with healthy individuals and/or with other CF-sharing features disorders will reveal a solid contribution to provide additional information on CF pathogenesis that ultimately have to be translated in the identification of diagnostic/prognostic biomarkers and therapeutics molecular targets for drug therapies [266].

8.1 The Fate of Proteomics in CF: Biomarkers Identification

In CF, biomarkers can be useful in diagnosis confirmation and/or evaluation of disease progression or patient response to therapeutic interventions. Biomarkers as predictive indicators of early events of CF pulmonary exacerbations can also be applyied as starting point for more aggressive treatments aimed to limit the extension of lung damage. However, biomarkers for CF, is still at an early stage of discovery, and therefore a research area of considerable interest to the Clinical Proteomics. Several concerns must be addressed before conducting biomarker studies in CF.

Numerous biological samples and proteomic-based approaches can be chosen depending on the type of question addressed (**Figure I.8**).

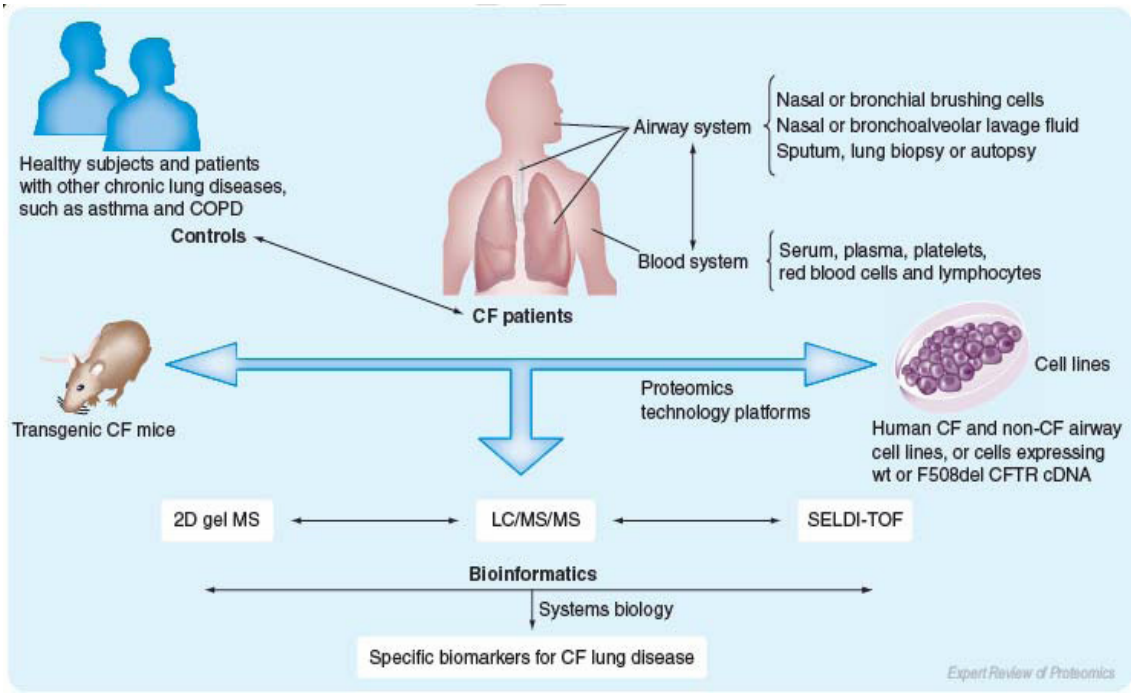


Figure I.8: Schematic overview on examples of samples and strategies used by Proteomics to address the biological question of biomarkers discovery for CF lung disease (Adapted from [266]).

Some recent works have unravel important information on the proteinaceous environment in CF by making use of cell lines [60,212,267-270], transgenic mice [271-273] and various types of human samples (see next references). In spite of the use of these biological samples, human lung tissue would be the ideal biological sample to work with in biomarkers discovery when dealing with a respiratory disease as CF. However, samples from lung transplantation present a high degree of differentiation and degradation, making them unrepresentative of a patient's native lung tissues. Also, collecting lung tissue for investigation studies out of the medical context can raise ethical constrains. One alternative to overcome this limitation and be able to profile cell and protein content of both upper and lower parts of the respiratory tract is to use representative samples of such portions: sputum [224,274-276], mucus [134,277], bronchoalveolar lavage fluid (BALF) [232,278,279], airway surface liquid (ASL) [131,280] nasal epithelial cells [281,282] or bronchial brushing cells [283]. Studies on such human biological samples have reported sets of proteins being differentially

expressed in CF when compared with non-CF individuals or with those with similar lung pathologies, such as asthma or chronic obstructive pulmonary disease (COPD).

Once discvovered, the candidate protein biomarkers should be validated and translated into clinical assays, preferentially measurable in samples obtained by rapid and non-invasive methods making serum one of the most promising samples to analyze [284,285]. Two different complementary approaches have already provided important results by screening the serum of CF patients. Srivastava *et al.* reported a CF-specific serum proteomics signature comprising proteins mediators of inflammation from the NFκB signalling pathway [167] while Pedersen and co-workers identified autoantibodies in serum of CF patients against inflammatory and defense proteins by an immunoproteomic approach [275]. The development of protein profiling techniques will be able to outline risk factors for susceptible patients, predict how well a patient will respond to a particular treatment and monitor improvements made during the course of a therapy [286].

In brief, proteomics will continue to be a key point in CF investigation bringing new insights to the elucidation of CF basic mechanisms that will culminate in the validation of protein biomarkers with potential clinical interest and translation into effective therapies for lung disease.

REFERENCES

[1] Welsh MJ, Smith AE. Cystic fibrosis. Sci Am 1995; 273:52-9.

[2] Busch R. On the history of cystic fibrosis. Acta Univ Carol Med (Praha) 1990; 36:13-5.

[3] Fanconi G. Das coeliakiesyndrom bei angeborener zysticher pankreasfibromatose und bronchiektasien. Wien Med Wschr 1936; 86:753-756.

[4] Anderson DH. Cystic Fibrosis of the pancreas and its relation to celiac disease: a clinical and pathological study. Am J Dis Child 1938; 56:344-399.

[5] Farber S. Some organic digestive disturbances in early life. J Mich Med Soc 1945; 44:587.

[6] Di Sant'Agnese PA, Darling RC, Perera GA, Shea E. Abnormal electrolyte composition of sweat in cystic fibrosis of the pancreas; clinical significance and relationship to the disease. Pediatrics 1953; 12:549-63.

[7] Gibson LE, Cooke RE. A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis. Pediatrics 1959; 23:545-9.

[8] Knowles M, Gatzy J, Boucher R. Increased bioelectric potential difference across respiratory epithelia in cystic fibrosis. N Engl J Med 1981; 305:1489-95.

[9] Quinton PM. Chloride impermeability in cystic fibrosis. Nature 1983; 301:421-2.

[10] Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, *et al.* Identification of the cystic fibrosis gene: genetic analysis. Science 1989; 245:1073-80.
[11] Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, *et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989; 245:1066-73.
[12] Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G, *et al.* Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 1989; 245:1059-65.
[13] Orenstein D, Rosenstein BJ, Stern R. Diagnosis of Cystic Fibrosis. in: D. Orenstein, B.J. Rosenstein, and R. Stern, (Eds.), Cystic Fibrosis medical care, Lippincott Williams&Wilkins, Philadelphia, 2000, pp. 21-54.
[14] Collins FS. Cystic fibrosis: molecular biology and therapeutic implications. Science 1992; 256:774-9.
[15] Cystic Fibrosis Foundation. Cystic Fibrosis Foundation, Patient Registry 2005 Annual Report, Bethesda, MD. 2008.
[16] Jilling T, Kirk KL. The biogenesis, traffic, and function of the cystic fibrosis transmembrane conductance regulator. Int Rev Cytol 1997; 172:193-241.
[17] Davis PB, Drumm M, Konstan MW. Cystic fibrosis. Am J Respir Crit Care Med 1996; 154:1229-56.
[18] Padman R, McColley SA, Miller DP, Konstan MW, Morgan WJ, *et al.* Infant care patterns at epidemiologic study of cystic fibrosis sites that achieve superior childhood lung function. Pediatrics 2007; 119:e531-7.
[19] Rosenstein BJ, Cutting GR. The diagnosis of cystic fibrosis: a consensus statement. Cystic Fibrosis Foundation Consensus Panel. J Pediatr 1998; 132:589-95.
[20] Ratjen F, Doring G. Cystic fibrosis. Lancet 2003; 361:681-9.
[21] Strausbaugh SD, Davis PB. Cystic fibrosis: a review of epidemiology and pathobiology. Clin Chest Med 2007; 28:279-88.
[22] Assael BM, Castellani C, Ocampo MB, Iansa P, Callegaro A, *et al.* Epidemiology and survival analysis of cystic fibrosis in an area of intense neonatal screening over 30 years. Am J Epidemiol 2002; 156:397-401.
[23] Quinton PM, Bijman J. Higher bioelectric potentials due to decreased chloride absorption in the sweat glands of patients with cystic fibrosis. N Engl J Med 1983; 308:1185-9.
[24] Stern RC. The diagnosis of cystic fibrosis. N Engl J Med 1997; 336:487-91.
[25] White NL, Higgins CF, Trezise AE. Tissue-specific in vivo transcription start sites of the human and murine cystic fibrosis genes. Hum Mol Genet 1998; 7:363-9.
[26] Zielenski J. Genotype and phenotype in cystic fibrosis. Respiration 2000; 67:117-33.
[27] Riordan JR. The cystic fibrosis transmembrane conductance regulator. Annu Rev Physiol 1993; 55:609-30.
[28] Chu CS, Trapnell BC, Curristin SM, Cutting GR, Crystal RG. Extensive posttranscriptional deletion of the coding sequences for part of nucleotide-binding fold 1 in respiratory epithelial mRNA transcripts of the cystic fibrosis transmembrane conductance regulator gene is not associated with the clinical manifestations of cystic fibrosis. J Clin Invest 1992; 90:785-90.
[29] Strong TV, Wilkinson DJ, Mansoura MK, Devor DC, Henze K, *et al.* Expression of an abundant alternatively spliced form of the cystic fibrosis transmembrane conductance regulator (CFTR) gene is not associated with a cAMP-activated chloride conductance. Hum Mol Genet 1993; 2:225-30.

[30] Sheppard DN, Welsh MJ. Structure and function of the CFTR chloride channel. Physiol Rev 1999; 79:S23-45.
[31] Akabas MH. Cystic fibrosis transmembrane conductance regulator. Structure and function of an epithelial chloride channel. J Biol Chem 2000; 275:3729-32.
[32] Schillers H, Shahin V, Albermann L, Schafer C, Oberleithner H. Imaging CFTR: a tail to tail dimer with a central pore. Cell Physiol Biochem 2004; 14:1-10.
[33] Rosenberg MF, Kamis AB, Aleksandrov LA, Ford RC, Riordan JR. Purification and crystallization of the cystic fibrosis transmembrane conductance regulator (CFTR). J Biol Chem 2004; 279:39051-7.
[34] Serohijos AW, Hegedus T, Aleksandrov AA, He L, Cui L, *et al.* Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D structure crucial to assembly and channel function. Proc Natl Acad Sci U S A 2008; 105:3256-61.
[35] Serohijos AW, Hegedus T, Riordan JR, Dokholyan NV. Diminished self-chaperoning activity of the DeltaF508 mutant of CFTR results in protein misfolding. PLoS Comput Biol 2008; 4:e1000008.
[36] Hughes AL. Evolution of the ATP-binding-cassette transmembrane transporters of vertebrates. Mol Biol Evol 1994; 11:899-910.
[37] Carson MR, Travis SM, Welsh MJ. The two nucleotide-binding domains of cystic fibrosis transmembrane conductance regulator (CFTR) have distinct functions in controlling channel activity. J Biol Chem 1995; 270:1711-7.
[38] Engelhardt JF, Yankaskas JR, Ernst SA, Yang Y, Marino CR, *et al.* Submucosal glands are the predominant site of CFTR expression in the human bronchus. Nat Genet 1992; 2:240-8.
[39] Trezise AE, Buchwald M. In vivo cell-specific expression of the cystic fibrosis transmembrane conductance regulator. Nature 1991; 353:434-7.
[40] Kartner N, Augustinas O, Jensen TJ, Naismith AL, Riordan JR. Mislocalization of delta F508 CFTR in cystic fibrosis sweat gland. Nat Genet 1992; 1:321-7.
[41] Trezise AE, Chambers JA, Wardle CJ, Gould S, Harris A. Expression of the cystic fibrosis gene in human foetal tissues. Hum Mol Genet 1993; 2:213-8.
[42] Richards B, Skoletsky J, Shuber AP, Balfour R, Stern RC, *et al.* Multiplex PCR amplification from the CFTR gene using DNA prepared from buccal brushes/swabs. Hum Mol Genet 1993; 2:159-63.
[43] Manson AL, Trezise AE, MacVinish LJ, Kasschau KD, Birchall N, *et al.* Complementation of null CF mice with a human CFTR YAC transgene. Embo J 1997; 16:4238-49.
[44] Yang Y, Raper SE, Cohn JA, Engelhardt JF, Wilson JM. An approach for treating the hepatobiliary disease of cystic fibrosis by somatic gene transfer. Proc Natl Acad Sci U S A 1993; 90:4601-5.
[45] Todd-Turla KM, Rusvai E, Naray-Fejes-Toth A, Fejes-Toth G. CFTR expression in cortical collecting duct cells. Am J Physiol 1996; 270:F237-44.
[46] Mulberg AE, Weyler RT, Altschuler SM, Hyde TM. Cystic fibrosis transmembrane conductance regulator expression in human hypothalamus. Neuroreport 1998; 9:141-4.
[47] Robert R, Thoreau V, Norez C, Cantereau A, Kitzis A, *et al.* Regulation of the cystic fibrosis transmembrane conductance regulator channel by beta-adrenergic agonists and vasoactive intestinal peptide in rat smooth muscle cells and its role in vasorelaxation. J Biol Chem 2004; 279:21160-8.

[48] Robert R, Norez C, Becq F. Disruption of CFTR chloride channel alters mechanical properties and cAMP-dependent Cl- transport of mouse aortic smooth muscle cells. J Physiol 2005; 568:483-95.
[49] Nagel G, Hwang TC, Nastiuk KL, Nairn AC, Gadsby DC. The protein kinase A-regulated cardiac Cl- channel resembles the cystic fibrosis transmembrane conductance regulator. Nature 1992; 360:81-4.
[50] Kulaksiz H, Schlenker T, Rost D, Stiehl A, Volkmann M*, et al.* Guanylin regulates chloride secretion in the human gallbladder via the bile fluid. Gastroenterology 2004; 126:732-40.
[51] Yoshimura K, Nakamura H, Trapnell BC, Chu CS, Dalemans W*, et al.* Expression of the cystic fibrosis transmembrane conductance regulator gene in cells of non-epithelial origin. Nucleic Acids Res 1991; 19:5417-23.
[52] Sprague RS, Ellsworth ML, Stephenson AH, Kleinhenz ME, Lonigro AJ. Deformation-induced ATP release from red blood cells requires CFTR activity. Am J Physiol 1998; 275:H1726-32.
[53] Lange T, Jungmann P, Haberle J, Falk S, Duebbers A*, et al.* Reduced number of CFTR molecules in erythrocyte plasma membrane of cystic fibrosis patients. Mol Membr Biol 2006; 23:317-23.
[54] McCray PB, Jr., Reenstra WW, Louie E, Johnson J, Bettencourt JD*, et al.* Expression of CFTR and presence of cAMP-mediated fluid secretion in human fetal lung. Am J Physiol 1992; 262:L472-81.
[55] Engelhardt JF, Zepeda M, Cohn JA, Yankaskas JR, Wilson JM. Expression of the cystic fibrosis gene in adult human lung. J Clin Invest 1994; 93:737-49.
[56] Beck S, Penque D, Garcia S, Gomes A, Farinha C*, et al.* Cystic fibrosis patients with the 3272-26A-->G mutation have mild disease, leaky alternative mRNA splicing, and CFTR protein at the cell membrane. Hum Mutat 1999; 14:133-44.
[57] Puchelle E, Gaillard D, Ploton D, Hinnrasky J, Fuchey C*, et al.* Differential localization of the cystic fibrosis transmembrane conductance regulator in normal and cystic fibrosis airway epithelium. Am J Respir Cell Mol Biol 1992; 7:485-91.
[58] Kalin N, Claass A, Sommer M, Puchelle E, Tummler B. DeltaF508 CFTR protein expression in tissues from patients with cystic fibrosis. J Clin Invest 1999; 103:1379-89.
[59] Wang X, Koulov AV, Kellner WA, Riordan JR, Balch WE. Chemical and biological folding contribute to temperature-sensitive DeltaF508 CFTR trafficking. Traffic 2008; 9:1878-93.
[60] Gomes-Alves P, Couto F, Pesquita C, Coelho AV, Penque D. Rescue of F508del-CFTR by RXR motif inactivation triggers proteome modulation associated with the unfolded protein response. Biochim Biophys Acta 2010; 1804:856-865.
[61] Cholon DM, O'Neal WK, Randell SH, Riordan JR, Gentzsch M. Modulation of endocytic trafficking and apical stability of CFTR in primary human airway epithelial cultures. Am J Physiol Lung Cell Mol Physiol 2010; 298:L304-14.
[62] Drumm ML, Pope HA, Cliff WH, Rommens JM, Marvin SA*, et al.* Correction of the cystic fibrosis defect in vitro by retrovirus-mediated gene transfer. Cell 1990; 62:1227-33.
[63] Rich DP, Anderson MP, Gregory RJ, Cheng SH, Paul S*, et al.* Expression of cystic fibrosis transmembrane conductance regulator corrects defective chloride channel regulation in cystic fibrosis airway epithelial cells. Nature 1990; 347:358-63.

[64] Gregory RJ, Cheng SH, Rich DP, Marshall J, Paul S*, et al.* Expression and characterization of the cystic fibrosis transmembrane conductance regulator. Nature 1990; 347:382-6.
[65] Bear CE, Li CH, Kartner N, Bridges RJ, Jensen TJ*, et al.* Purification and functional reconstitution of the cystic fibrosis transmembrane conductance regulator (CFTR). Cell 1992; 68:809-18.
[66] Vergani P, Basso C, Mense M, Nairn AC, Gadsby DC. Control of the CFTR channel's gates. Biochem Soc Trans 2005; 33:1003-7.
[67] Vergani P, Lockless SW, Nairn AC, Gadsby DC. CFTR channel opening by ATP-driven tight dimerization of its nucleotide-binding domains. Nature 2005; 433:876-80.
[68] Boucher RC, Stutts MJ, Knowles MR, Cantley L, Gatzy JT. Na+ transport in cystic fibrosis respiratory epithelia. Abnormal basal rate and response to adenylate cyclase activation. J Clin Invest 1986; 78:1245-52.
[69] Stutts MJ, Canessa CM, Olsen JC, Hamrick M, Cohn JA*, et al.* CFTR as a cAMP-dependent regulator of sodium channels. Science 1995; 269:847-50.
[70] Wirbel RJ, Schulte M, Mutschler W. [Primary malignant bone and soft tissue tumors of the pelvis--oncologic and functional outcome]. Langenbecks Arch Chir Suppl Kongressbd 1996; 113:914-7.
[71] Wei L, Vankeerberghen A, Cuppens H, Eggermont J, Cassiman JJ*, et al.* Interaction between calcium-activated chloride channels and the cystic fibrosis transmembrane conductance regulator. Pflugers Arch 1999; 438:635-41.
[72] Jovov B, Ismailov, II, Berdiev BK, Fuller CM, Sorscher EJ*, et al.* Interaction between cystic fibrosis transmembrane conductance regulator and outwardly rectified chloride channels. J Biol Chem 1995; 270:29194-200.
[73] Jovov B, Ismailov, II, Benos DJ. Cystic fibrosis transmembrane conductance regulator is required for protein kinase A activation of an outwardly rectified anion channel purified from bovine tracheal epithelia. J Biol Chem 1995; 270:1521-8.
[74] McNicholas CM, Guggino WB, Schwiebert EM, Hebert SC, Giebisch G*, et al.* Sensitivity of a renal K+ channel (ROMK2) to the inhibitory sulfonylurea compound glibenclamide is enhanced by coexpression with the ATP-binding cassette transporter cystic fibrosis transmembrane regulator. Proc Natl Acad Sci U S A 1996; 93:8083-8.
[75] Mall M, Wissner A, Schreiber R, Kuehr J, Seydewitz HH*, et al.* Role of K(V)LQT1 in cyclic adenosine monophosphate-mediated Cl(-) secretion in human airway epithelia. Am J Respir Cell Mol Biol 2000; 23:283-9.
[76] Hasegawa H, Skach W, Baker O, Calayag MC, Lingappa V*, et al.* A multifunctional aqueous channel formed by CFTR. Science 1992; 258:1477-9.
[77] Schreiber R, Greger R, Nitschke R, Kunzelmann K. Cystic fibrosis transmembrane conductance regulator activates water conductance in Xenopus oocytes. Pflugers Arch 1997; 434:841-7.
[78] Schreiber R, Nitschke R, Greger R, Kunzelmann K. The cystic fibrosis transmembrane conductance regulator activates aquaporin 3 in airway epithelial cells. J Biol Chem 1999; 274:11811-6.
[79] Gray MA, O'Reilly C, Winpenny J, Argent B. Functional interactions of HCO3- with cystic fibrosis transmembrane conductance regulator. Jop 2001; 2:207-11.
[80] Choi JY, Lee MG, Ko S, Muallem S. Cl(-)-dependent HCO3- transport by cystic fibrosis transmembrane conductance regulator. Jop 2001; 2:243-6.

[81] Choi JY, Muallem D, Kiselyov K, Lee MG, Thomas PJ, *et al.* Aberrant CFTR-dependent HCO3- transport in mutations associated with cystic fibrosis. Nature 2001; 410:94-7.
[82] Kogan I, Ramjeesingh M, Li C, Kidd JF, Wang Y, *et al.* CFTR directly mediates nucleotide-regulated glutathione flux. Embo J 2003; 22:1981-9.
[83] Roum JH, Buhl R, McElvaney NG, Borok Z, Crystal RG. Systemic deficiency of glutathione in cystic fibrosis. J Appl Physiol 1993; 75:2419-24.
[84] Wine JJ. Rules of conduct for the cystic fibrosis anion channel. Nat Med 2003; 9:827-8.
[85] Bradbury NA, Clark JA, Watkins SC, Widnell CC, Smith HSt, *et al.* Characterization of the internalization pathways for the cystic fibrosis transmembrane conductance regulator. Am J Physiol 1999; 276:L659-68.
[86] Saiman L, Prince A. Pseudomonas aeruginosa pili bind to asialoGM1 which is increased on the surface of cystic fibrosis epithelial cells. J Clin Invest 1993; 92:1875-80.
[87] Kowalski MP, Pier GB. Localization of cystic fibrosis transmembrane conductance regulator to lipid rafts of epithelial cells is required for Pseudomonas aeruginosa-induced cellular activation. J Immunol 2004; 172:418-25.
[88] Pier GB, Grout M, Zaidi TS. Cystic fibrosis transmembrane conductance regulator is an epithelial cell receptor for clearance of Pseudomonas aeruginosa from the lung. Proc Natl Acad Sci U S A 1997; 94:12088-93.
[89] Glick MC, Kothari VA, Liu A, Stoykova LI, Scanlin TF. Activity of fucosyltransferases and altered glycosylation in cystic fibrosis airway epithelial cells. Biochimie 2001; 83:743-7.
[90] Prince LS, Workman RB, Jr., Marchase RB. Rapid endocytosis of the cystic fibrosis transmembrane conductance regulator chloride channel. Proc Natl Acad Sci U S A 1994; 91:5192-6.
[91] Bertrand CA, Frizzell RA. The role of regulated CFTR trafficking in epithelial secretion. Am J Physiol Cell Physiol 2003; 285:C1-18.
[92] Zielenski J, Tsui LC. Cystic fibrosis: genotypic and phenotypic variations. Annu Rev Genet 1995; 29:777-807.
[93] The Cystic Fibrosis Genotype-Phenotype Consortium. Correlation between genotype and phenotype in patients with cystic fibrosis. The Cystic Fibrosis Genotype-Phenotype Consortium. N Engl J Med 1993; 329:1308-13.
[94] Vankeerberghen A, Cuppens H, Cassiman JJ. The cystic fibrosis transmembrane conductance regulator: an intriguing protein with pleiotropic functions. J Cyst Fibros 2002; 1:13-29.
[95] Haardt M, Benharouga M, Lechardeur D, Kartner N, Lukacs GL. C-terminal truncations destabilize the cystic fibrosis transmembrane conductance regulator without impairing its biogenesis. A novel class of mutation. J Biol Chem 1999; 274:21873-7.
[96] Noone PG, Knowles MR. 'CFTR-opathies': disease phenotypes associated with cystic fibrosis transmembrane regulator gene mutations. Respir Res 2001; 2:328-32.
[97] Cheng SH, Gregory RJ, Marshall J, Paul S, Souza DW, *et al.* Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. Cell 1990; 63:827-34.
[98] Dalemans W, Barbry P, Champigny G, Jallat S, Dott K, *et al.* Altered chloride ion channel kinetics associated with the delta F508 cystic fibrosis mutation. Nature 1991; 354:526-8.

[99] Lukacs GL, Mohamed A, Kartner N, Chang XB, Riordan JR*, et al.* Conformational maturation of CFTR but not its mutant counterpart (delta F508) occurs in the endoplasmic reticulum and requires ATP. Embo J 1994; 13:6076-86.
[100] Wang L, Freedman SD. Laboratory tests for the diagnosis of cystic fibrosis. Am J Clin Pathol 2002; 117 Suppl:S109-15.
[101] Baumer JH. Evidence based guidelines for the performance of the sweat test for the investigation of cystic fibrosis in the UK. Arch Dis Child 2003; 88:1126-7.
[102] LeGrys VA. Sweat testing for the diagnosis of cystic fibrosis: practical considerations. J Pediatr 1996; 129:892-7.
[103] Farrell PM, Koscik RE. Sweat chloride concentrations in infants homozygous or heterozygous for F508 cystic fibrosis. Pediatrics 1996; 97:524-8.
[104] Muller F, Simon-Bouy B, Girodon E, Monnier N, Malinge MC*, et al.* Predicting the risk of cystic fibrosis with abnormal ultrasound signs of fetal bowel: results of a French molecular collaborative study based on 641 prospective cases. Am J Med Genet 2002; 110:109-15.
[105] Hardcastle J, Hardcastle PT, Taylor CJ, Goldhill J. Failure of cholinergic stimulation to induce a secretory response from the rectal mucosa in cystic fibrosis. Gut 1991; 32:1035-9.
[106] Vawter GF, Shwachman H. Cystic fibrosis in adults: an autopsy study. Pathol Annu 1979; 14 Pt 2:357-82.
[107] Bhaskar KR, Turner BS, Grubman SA, Jefferson DM, LaMont JT. Dysregulation of proteoglycan production by intrahepatic biliary epithelial cells bearing defective (delta-f508) cystic fibrosis transmembrane conductance regulator. Hepatology 1998; 27:7-14.
[108] Lanng S, Thorsteinsson B, Nerup J, Koch C. Diabetes mellitus in cystic fibrosis: effect of insulin therapy on lung function and infections. Acta Paediatr 1994; 83:849-53.
[109] Dodge JA. Male fertility in cystic fibrosis. Lancet 1995; 346:587-8.
[110] Davis PB. Cystic fibrosis since 1938. Am J Respir Crit Care Med 2006; 173:475-82.
[111] Johannesson M, Landgren BM, Csemiczky G, Hjelte L, Gottlieb C. Female patients with cystic fibrosis suffer from reproductive endocrinological disorders despite good clinical status. Hum Reprod 1998; 13:2092-7.
[112] Johannesson M, Carlson M, Brucefors AB, Hjelte L. Cystic fibrosis through a female perspective: psychosocial issues and information concerning puberty and motherhood. Patient Educ Couns 1998; 34:115-23.
[113] Jarzabek K, Zbucka M, Pepinski W, Szamatowicz J, Domitrz J*, et al.* Cystic fibrosis as a cause of infertility. Reprod Biol 2004; 4:119-29.
[114] Cystic Fibrosis Genotype-Phenotype Consortium. Correlation between genotype and phenotype in patients with cystic fibrosis. N Engl J Med 1993; 329:1308-13.
[115] Culver CM, Clouser KD, Gert B, Brody H, Fletcher J*, et al.* Basic curricular goals in medical ethics. N Engl J Med 1985; 312:253-6.
[116] Holl RW, Wolf A, Thon A, Bernhard M, Buck C*, et al.* Insulin resistance with altered secretory kinetics and reduced proinsulin in cystic fibrosis patients. J Pediatr Gastroenterol Nutr 1997; 25:188-93.
[117] Baker SS, Borowitz D, Baker RD. Pancreatic exocrine function in patients with cystic fibrosis. Curr Gastroenterol Rep 2005; 7:227-33.
[118] Bals R, Weiner DJ, Wilson JM. The innate immune system in cystic fibrosis lung disease. J Clin Invest 1999; 103:303-7.

[119] Wanner A, Salathe M, O'Riordan TG. Mucociliary clearance in the airways. Am J Respir Crit Care Med 1996; 154:1868-902.
[120] Goldman MJ, Anderson GM, Stolzenberg ED, Kari UP, Zasloff M*, et al.* Human beta-defensin-1 is a salt-sensitive antibiotic in lung that is inactivated in cystic fibrosis. Cell 1997; 88:553-60.
[121] Coakley RD, Boucher RC. Regulation and functional significance of airway surface liquid pH. Jop 2001; 2:294-300.
[122] Worlitzsch D, Tarran R, Ulrich M, Schwab U, Cekici A*, et al.* Effects of reduced mucus oxygen concentration in airway Pseudomonas infections of cystic fibrosis patients. J Clin Invest 2002; 109:317-25.
[123] Pier GB, Grout M, Zaidi TS, Olsen JC, Johnson LG*, et al.* Role of mutant CFTR in hypersusceptibility of cystic fibrosis patients to lung infections. Science 1996; 271:64-7.
[124] Boucher RC. Molecular insights into the physiology of the 'thin film' of airway surface liquid. J Physiol 1999; 516 (Pt 3):631-8.
[125] Verkman AS, Song Y, Thiagarajah JR. Role of airway surface liquid and submucosal glands in cystic fibrosis lung disease. Am J Physiol Cell Physiol 2003; 284:C2-15.
[126] Imundo L, Barasch J, Prince A, Al-Awqati Q. Cystic fibrosis epithelial cells have a receptor for pathogenic bacteria on their apical surface. Proc Natl Acad Sci U S A 1995; 92:3019-23.
[127] Zuelzer WW, Newton WA, Jr. The pathogenesis of fibrocystic disease of the pancreas; a study of 36 cases with special reference to the pulmonary lesions. Pediatrics 1949; 4:53-69.
[128] Mall M, Boucher RC. Pathogenesis of Pulmonary Disease in Cystic Fibrosis. in: A. Bush, E.W.F.W. Alton, J.C. Davies, and U. Griesenbach, (Eds.), Cystic Fibrosis in the 21st Century, Karger, Basel, 2006, pp. 116-121.
[129] Knowles MR, Boucher RC. Mucus clearance as a primary innate defense mechanism for mammalian airways. J Clin Invest 2002; 109:571-7.
[130] Matsui H, Davis CW, Tarran R, Boucher RC. Osmotic water permeabilities of cultured, well-differentiated normal and cystic fibrosis airway epithelia. J Clin Invest 2000; 105:1419-27.
[131] Matsui H, Grubb BR, Tarran R, Randell SH, Gatzy JT*, et al.* Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease. Cell 1998; 95:1005-15.
[132] Tarran R, Grubb BR, Parsons D, Picher M, Hirsh AJ*, et al.* The CF salt controversy: in vivo observations and therapeutic approaches. Mol Cell 2001; 8:149-58.
[133] Tarran R, Grubb BR, Gatzy JT, Davis CW, Boucher RC. The relative roles of passive surface forces and active ion transport in the modulation of airway surface liquid volume and composition. J Gen Physiol 2001; 118:223-36.
[134] Inglis SK, Wilson SM. Cystic fibrosis and airway submucosal glands. Pediatr Pulmonol 2005; 40:279-84.
[135] Rosenfeld M, Gibson RL, McNamara S, Emerson J, Burns JL*, et al.* Early pulmonary infection, inflammation, and clinical outcomes in infants with cystic fibrosis. Pediatr Pulmonol 2001; 32:356-66.
[136] Armstrong DS, Grimwood K, Carlin JB, Carzino R, Gutierrez JP*, et al.* Lower airway inflammation in infants and young children with cystic fibrosis. Am J Respir Crit Care Med 1997; 156:1197-204.

[137] Henry RL, Mellis CM, Petrovic L. Mucoid Pseudomonas aeruginosa is a marker of poor survival in cystic fibrosis. Pediatr Pulmonol 1992; 12:158-61.
[138] Frederiksen B, Koch C, Hoiby N. Changing epidemiology of Pseudomonas aeruginosa infection in Danish cystic fibrosis patients (1974-1995). Pediatr Pulmonol 1999; 28:159-66.
[139] Emerson J, Rosenfeld M, McNamara S, Ramsey B, Gibson RL. Pseudomonas aeruginosa and other predictors of mortality and morbidity in young children with cystic fibrosis. Pediatr Pulmonol 2002; 34:91-100.
[140] Khan TZ, Wagener JS, Bost T, Martinez J, Accurso FJ*, et al.* Early pulmonary inflammation in infants with cystic fibrosis. Am J Respir Crit Care Med 1995; 151:1075-82.
[141] Poschet JF, Boucher JC, Tatterson L, Skidmore J, Van Dyke RW*, et al.* Molecular basis for defective glycosylation and Pseudomonas pathogenesis in cystic fibrosis lung. Proc Natl Acad Sci U S A 2001; 98:13972-7.
[142] Smith JJ, Travis SM, Greenberg EP, Welsh MJ. Cystic fibrosis airway epithelia fail to kill bacteria because of abnormal airway surface fluid. Cell 1996; 85:229-36.
[143] Jiang C, Finkbeiner WE, Widdicombe JH, McCray PB, Jr., Miller SS. Altered fluid transport across airway epithelium in cystic fibrosis. Science 1993; 262:424-7.
[144] Ge Y, Ezzell RM, Warren HS. Localization of endotoxin in the rat intestinal epithelium. J Infect Dis 2000; 182:873-81.
[145] Hata JS, Fick RB, Jr. Airway adherence of Pseudomonas aeruginosa: mucoexopolysaccharide binding to human and bovine airway proteins. J Lab Clin Med 1991; 117:410-22.
[146] Oliver A, Canton R, Campo P, Baquero F, Blazquez J. High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection. Science 2000; 288:1251-4.
[147] Beringer PM, Appleman MD. Unusual respiratory bacterial flora in cystic fibrosis: microbiologic and clinical features. Curr Opin Pulm Med 2000; 6:545-50.
[148] Hoiby N. Prospects for the prevention and control of pseudomonal infection in children with cystic fibrosis. Paediatr Drugs 2000; 2:451-63.
[149] Stutman HR, Lieberman JM, Nussbaum E, Marks MI. Antibiotic prophylaxis in infants and young children with cystic fibrosis: a randomized controlled trial. J Pediatr 2002; 140:299-305.
[150] Solis A, Brown D, Hughes J, Van Saene HK, Heaf DP. Methicillin-resistant Staphylococcus aureus in children with cystic fibrosis: An eradication protocol. Pediatr Pulmonol 2003; 36:189-95.
[151] Burkholder W. Sour skin, a bacterial rot of onion bulbs. Phytopathology 1950; 40:115-117.
[152] Govan JR, Hughes JE, Vandamme P. Burkholderia cepacia: medical, taxonomic and ecological issues. J Med Microbiol 1996; 45:395-407.
[153] LiPuma JJ. Burkholderia cepacia. Management issues and new insights. Clin Chest Med 1998; 19:473-86, vi.
[154] Mahenthiralingam E, Baldwin A, Vandamme P. Burkholderia cepacia complex infection in patients with cystic fibrosis. J Med Microbiol 2002; 51:533-8.
[155] Aris RM, Routh JC, LiPuma JJ, Heath DG, Gilligan PH. Lung transplantation for cystic fibrosis patients with Burkholderia cepacia complex. Survival linked to genomovar type. Am J Respir Crit Care Med 2001; 164:2102-6.

[156] De Soyza A, Morris K, McDowell A, Doherty C, Archer L*, et al.* Prevalence and clonality of Burkholderia cepacia complex genomovars in UK patients with cystic fibrosis referred for lung transplantation. Thorax 2004; 59:526-8.
[157] Aaron SD, Ferris W, Henry DA, Speert DP, Macdonald NE. Multiple combination bactericidal antibiotic testing for patients with cystic fibrosis infected with Burkholderia cepacia. Am J Respir Crit Care Med 2000; 161:1206-12.
[158] Zahm JM, Gaillard D, Dupuit F, Hinnrasky J, Porteous D*, et al.* Early alterations in airway mucociliary clearance and inflammation of the lamina propria in CF mice. Am J Physiol 1997; 272:C853-9.
[159] Ornoy A, Arnon J, Katznelson D, Granat M, Caspi B*, et al.* Pathological confirmation of cystic fibrosis in the fetus following prenatal diagnosis. Am J Med Genet 1987; 28:935-47.
[160] Blackwell TS, Stecenko AA, Christman JW. Dysregulated NF-kappaB activation in cystic fibrosis: evidence for a primary inflammatory disorder. Am J Physiol Lung Cell Mol Physiol 2001; 281:L69-70.
[161] Joseph T, Look D, Ferkol T. NF-kappaB activation and sustained IL-8 gene expression in primary cultures of cystic fibrosis airway epithelial cells stimulated with Pseudomonas aeruginosa. Am J Physiol Lung Cell Mol Physiol 2005; 288:L471-9.
[162] Bonfield TL, Konstan MW, Burfeind P, Panuska JR, Hilliard JB*, et al.* Normal bronchial epithelial cells constitutively produce the anti-inflammatory cytokine interleukin-10, which is downregulated in cystic fibrosis. Am J Respir Cell Mol Biol 1995; 13:257-61.
[163] Chmiel JF, Konstan MW, Knesebeck JE, Hilliard JB, Bonfield TL*, et al.* IL-10 attenuates excessive inflammation in chronic Pseudomonas infection in mice. Am J Respir Crit Care Med 1999; 160:2040-7.
[164] Moore KW, de Waal Malefyt R, Coffman RL, O'Garra A. Interleukin-10 and the interleukin-10 receptor. Annu Rev Immunol 2001; 19:683-765.
[165] Venkatakrishnan A, Stecenko AA, King G, Blackwell TR, Brigham KL*, et al.* Exaggerated activation of nuclear factor-kappaB and altered IkappaB-beta processing in cystic fibrosis bronchial epithelial cells. Am J Respir Cell Mol Biol 2000; 23:396-403.
[166] Weber AJ, Soong G, Bryan R, Saba S, Prince A. Activation of NF-kappaB in airway epithelial cells is dependent on CFTR trafficking and Cl- channel function. Am J Physiol Lung Cell Mol Physiol 2001; 281:L71-8.
[167] Srivastava M, Eidelman O, Jozwik C, Paweletz C, Huang W*, et al.* Serum proteomic signature for cystic fibrosis using an antibody microarray platform. Mol Genet Metab 2006; 87:303-10.
[168] van der Vliet A, Cross CE. Phagocyte oxidants and nitric oxide in cystic fibrosis: new therapeutic targets? Curr Opin Pulm Med 2000; 6:533-9.
[169] Schroeder TH, Lee MM, Yacono PW, Cannon CL, Gerceker AA*, et al.* CFTR is a pattern recognition molecule that extracts Pseudomonas aeruginosa LPS from the outer membrane into epithelial cells and activates NF-kappa B translocation. Proc Natl Acad Sci U S A 2002; 99:6907-12.
[170] Reiniger N, Lee MM, Coleman FT, Ray C, Golan DE*, et al.* Resistance to Pseudomonas aeruginosa chronic lung infection requires cystic fibrosis transmembrane conductance regulator-modulated interleukin-1 (IL-1) release and signaling through the IL-1 receptor. Infect Immun 2007; 75:1598-608.

[171] Dudez T, Borot F, Huang S, Kwak BR, Bacchetta M*, et al.* CFTR in a lipid raft-TNFR1 complex modulates gap junctional intercellular communication and IL-8 secretion. Biochim Biophys Acta 2008; 1783:779-88.
[172] Grassme H, Becker KA, Zhang Y, Gulbins E. Ceramide in bacterial infections and cystic fibrosis. Biol Chem 2008; 389:1371-9.
[173] Bodas M, Vij N. The NF-kappaB signaling in cystic fibrosis lung disease: pathophysiology and therapeutic potential. Discov Med 2010; 9:346-56.
[174] Tirouvanziam R, de Bentzmann S, Hubeau C, Hinnrasky J, Jacquot J*, et al.* Inflammation and infection in naive human cystic fibrosis airway grafts. Am J Respir Cell Mol Biol 2000; 23:121-7.
[175] Escotte S, Danel C, Gaillard D, Benoit S, Jacquot J*, et al.* Fluticasone propionate inhibits lipopolysaccharide-induced proinflammatory response in human cystic fibrosis airway grafts. J Pharmacol Exp Ther 2002; 302:1151-7.
[176] Cohn LA, Weber A, Phillips T, Lory S, Kaplan M*, et al.* Pseudomonas aeruginosa infection of respiratory epithelium in a cystic fibrosis xenograft model. J Infect Dis 2001; 183:919-27.
[177] Davies JC, Alton EW, Bush A. Cystic fibrosis. Bmj 2007; 335:1255-9.
[178] Konstan MW, Hilliard KA, Norvell TM, Berger M. Bronchoalveolar lavage findings in cystic fibrosis patients with stable, clinically mild lung disease suggest ongoing infection and inflammation. Am J Respir Crit Care Med 1994; 150:448-54.
[179] Tsui LC, Durie P. Genotype and phenotype in cystic fibrosis. Hosp Pract (Minneap) 1997; 32:115-8, 123-9, 134, passim.
[180] Kristidis P, Bozon D, Corey M, Markiewicz D, Rommens J*, et al.* Genetic determination of exocrine pancreatic function in cystic fibrosis. Am J Hum Genet 1992; 50:1178-84.
[181] Dray X, Kanaan R, Bienvenu T, Desmazes-Dufeu N, Dusser D*, et al.* Malnutrition in adults with cystic fibrosis. Eur J Clin Nutr 2005; 59:152-4.
[182] Patrizio P, Zielenski J. Congenital absence of the vas deferens: a mild form of cystic fibrosis. Mol Med Today 1996; 2:24-31.
[183] Schaedel C, de Monestrol I, Hjelte L, Johannesson M, Kornfalt R*, et al.* Predictors of deterioration of lung function in cystic fibrosis. Pediatr Pulmonol 2002; 33:483-91.
[184] Koch C, Cuppens H, Rainisio M, Madessani U, Harms H*, et al.* European Epidemiologic Registry of Cystic Fibrosis (ERCF): comparison of major disease manifestations between patients with different classes of mutations. Pediatr Pulmonol 2001; 31:1-12.
[185] Kneussl MP, Lang IM, Brenot FP. Medical management of primary pulmonary hypertension. Eur Respir J 1996; 9:2401-9.
[186] Johansen HK, Nir M, Hoiby N, Koch C, Schwartz M. Severity of cystic fibrosis in patients homozygous and heterozygous for delta F508 mutation. Lancet 1991; 337:631-4.
[187] Lai HJ, Cheng Y, Cho H, Kosorok MR, Farrell PM. Association between initial disease presentation, lung disease outcomes, and survival in patients with cystic fibrosis. Am J Epidemiol 2004; 159:537-46.
[188] Loubieres Y, Grenet D, Simon-Bouy B, Medioni J, Landais P*, et al.* Association between genetically determined pancreatic status and lung disease in adult cystic fibrosis patients. Chest 2002; 121:73-80.

[189] Rosenecker J. Relations between the frequency of the DeltaF 508 mutation and the course of pulmonary disease in cystic fibrosis patients infected with Pseudomonas aeruginosa. Eur J Med Res 2000; 5:356-9.
[190] Kerem E, Corey M, Kerem BS, Rommens J, Markiewicz D*, et al.* The relation between genotype and phenotype in cystic fibrosis--analysis of the most common mutation (delta F508). N Engl J Med 1990; 323:1517-22.
[191] Estivill X, Ortigosa L, Perez-Frias J, Dapena J, Ferrer J*, et al.* Clinical characteristics of 16 cystic fibrosis patients with the missense mutation R334W, a pancreatic insufficiency mutation with variable age of onset and interfamilial clinical differences. Hum Genet 1995; 95:331-6.
[192] Kiesewetter S, Macek M, Jr., Davis C, Curristin SM, Chu CS*, et al.* A mutation in CFTR produces different phenotypes depending on chromosomal background. Nat Genet 1993; 5:274-8.
[193] Monaghan KG, Highsmith WE, Amos J, Pratt VM, Roa B*, et al.* Genotype-phenotype correlation and frequency of the 3199del6 cystic fibrosis mutation among I148T carriers: results from a collaborative study. Genet Med 2004; 6:421-5.
[194] Ramalho AS, Beck S, Meyer M, Penque D, Cutting GR*, et al.* Five percent of normal cystic fibrosis transmembrane conductance regulator mRNA ameliorates the severity of pulmonary disease in cystic fibrosis. Am J Respir Cell Mol Biol 2002; 27:619-27.
[195] Collaco JM, Cutting GR. Update on gene modifiers in cystic fibrosis. Curr Opin Pulm Med 2008; 14:559-66.
[196] Vanscoy LL, Blackman SM, Collaco JM, Bowers A, Lai T*, et al.* Heritability of lung disease severity in cystic fibrosis. Am J Respir Crit Care Med 2007; 175:1036-43.
[197] Clarke LL, Grubb BR, Yankaskas JR, Cotton CU, McKenzie A*, et al.* Relationship of a non-cystic fibrosis transmembrane conductance regulator-mediated chloride conductance to organ-level disease in Cftr(-/-) mice. Proc Natl Acad Sci U S A 1994; 91:479-83.
[198] Rozmahel R, Wilschanski M, Matin A, Plyte S, Oliver M*, et al.* Modulation of disease severity in cystic fibrosis transmembrane conductance regulator deficient mice by a secondary genetic factor. Nat Genet 1996; 12:280-7.
[199] Zielenski J, Corey M, Rozmahel R, Markiewicz D, Aznarez I*, et al.* Detection of a cystic fibrosis modifier locus for meconium ileus on human chromosome 19q13. Nat Genet 1999; 22:128-9.
[200] Bronsveld I, Mekus F, Bijman J, Ballmann M, Greipel J*, et al.* Residual chloride secretion in intestinal tissue of deltaF508 homozygous twins and siblings with cystic fibrosis. The European CF Twin and Sibling Study Consortium. Gastroenterology 2000; 119:32-40.
[201] Mekus F, Ballmann M, Bronsveld I, Bijman J, Veeze H*, et al.* Categories of deltaF508 homozygous cystic fibrosis twin and sibling pairs with distinct phenotypic characteristics. Twin Res 2000; 3:277-93.
[202] Cutting GR. Modifier genes in Mendelian disorders: the example of cystic fibrosis. Ann N Y Acad Sci 2010; 1214:57-69.
[203] Griesenbach U. AEWFW, Davies J. CF Modifier Genes. in: C.T. Bolliger, (Ed.), Cystic Fibrosis in the 21st Centuty, Karger, Cape Town, 2006.
[204] Li C, Naren AP. CFTR chloride channel in the apical compartments: spatiotemporal coupling to its interacting partners. Integr Biol (Camb) 2010; 2:161-77.

[205] Ahner A, Brodsky JL. Checkpoints in ER-associated degradation: excuse me, which way to the proteasome? Trends Cell Biol 2004; 14:474-8.
[206] Brodsky JL. Chaperoning the maturation of the cystic fibrosis transmembrane conductance regulator. Am J Physiol Lung Cell Mol Physiol 2001; 281:L39-42.
[207] Pind S, Riordan JR, Williams DB. Participation of the endoplasmic reticulum chaperone calnexin (p88, IP90) in the biogenesis of the cystic fibrosis transmembrane conductance regulator. J Biol Chem 1994; 269:12784-8.
[208] Loo MA, Jensen TJ, Cui L, Hou Y, Chang XB, *et al.* Perturbation of Hsp90 interaction with nascent CFTR prevents its maturation and accelerates its degradation by the proteasome. Embo J 1998; 17:6879-87.
[209] Youker RT, Walsh P, Beilharz T, Lithgow T, Brodsky JL. Distinct roles for the Hsp40 and Hsp90 molecular chaperones during cystic fibrosis transmembrane conductance regulator degradation in yeast. Mol Biol Cell 2004; 15:4787-97.
[210] Meacham GC, Patterson C, Zhang W, Younger JM, Cyr DM. The Hsc70 co-chaperone CHIP targets immature CFTR for proteasomal degradation. Nat Cell Biol 2001; 3:100-5.
[211] Alberti S, Bohse K, Arndt V, Schmitz A, Hohfeld J. The cochaperone HspBP1 inhibits the CHIP ubiquitin ligase and stimulates the maturation of the cystic fibrosis transmembrane conductance regulator. Mol Biol Cell 2004; 15:4003-10.
[212] Wang X, Venable J, LaPointe P, Hutt DM, Koulov AV, *et al.* Hsp90 cochaperone Aha1 downregulation rescues misfolding of CFTR in cystic fibrosis. Cell 2006; 127:803-15.
[213] Nakamura Y, Tate L, Ertl RF, Kawamoto M, Mio T, *et al.* Bronchial epithelial cells regulate fibroblast proliferation. Am J Physiol 1995; 269:L377-87.
[214] Kawamoto M, Romberger DJ, Nakamura Y, Adachi Y, Tate L, *et al.* Modulation of fibroblast type I collagen and fibronectin production by bovine bronchial epithelial cells. Am J Respir Cell Mol Biol 1995; 12:425-33.
[215] Ignotz RA, Massague J. Transforming growth factor-beta stimulates the expression of fibronectin and collagen and their incorporation into the extracellular matrix. J Biol Chem 1986; 261:4337-45.
[216] Awad MR, El-Gamel A, Hasleton P, Turner DM, Sinnott PJ, *et al.* Genotypic variation in the transforming growth factor-beta1 gene: association with transforming growth factor-beta1 production, fibrotic lung disease, and graft fibrosis after lung transplantation. Transplantation 1998; 66:1014-20.
[217] El-Gamel A, Awad M, Sim E, Hasleton P, Yonan N, *et al.* Transforming growth factor-beta1 and lung allograft fibrosis. Eur J Cardiothorac Surg 1998; 13:424-30.
[218] Arkwright PD, Laurie S, Super M, Pravica V, Schwarz MJ, *et al.* TGF-beta(1) genotype and accelerated decline in lung function of patients with cystic fibrosis. Thorax 2000; 55:459-62.
[219] Bremer LA, Blackman SM, Vanscoy LL, McDougal KE, Bowers A, *et al.* Interaction between a novel TGFB1 haplotype and CFTR genotype is associated with improved lung function in cystic fibrosis. Hum Mol Genet 2008; 17:2228-37.
[220] Bartlett JR, Friedman KJ, Ling SC, Pace RG, Bell SC, *et al.* Genetic modifiers of liver disease in cystic fibrosis. Jama 2009; 302:1076-83.
[221] Pulleyn LJ, Newton R, Adcock IM, Barnes PJ. TGFbeta1 allele association with asthma severity. Hum Genet 2001; 109:623-7.

[222] Wu L, Chau J, Young RP, Pokorny V, Mills GD, *et al.* Transforming growth factor-beta1 genotype and susceptibility to chronic obstructive pulmonary disease. Thorax 2004; 59:126-9.
[223] Bonfield TL, Panuska JR, Konstan MW, Hilliard KA, Hilliard JB, *et al.* Inflammatory cytokines in cystic fibrosis lungs. Am J Respir Crit Care Med 1995; 152:2111-8.
[224] Greally P, Hussein MJ, Cook AJ, Sampson AP, Piper PJ, *et al.* Sputum tumour necrosis factor-alpha and leukotriene concentrations in cystic fibrosis. Arch Dis Child 1993; 68:389-92.
[225] Hull J, Thomson AH. Contribution of genetic factors other than CFTR to disease severity in cystic fibrosis. Thorax 1998; 53:1018-21.
[226] Arkwright PD, Pravica V, Geraghty PJ, Super M, Webb AK, *et al.* End-organ dysfunction in cystic fibrosis: association with angiotensin I converting enzyme and cytokine gene polymorphisms. Am J Respir Crit Care Med 2003; 167:384-9.
[227] Low TY, Lymnay PA, Davies JC, U. G, Shen N, *et al.* The effect of polymorphism in inflammatory mediators of clinical phenotype in CF. Pediatr Pulmonol 2003; 25:219.
[228] Texereau J, Marullo S, Hubert D, Coste J, Dusser DJ, *et al.* Nitric oxide synthase 1 as a potential modifier gene of decline in lung function in patients with cystic fibrosis. Thorax 2004; 59:156-8.
[229] Grasemann H, Knauer N, Buscher R, Hubner K, Drazen JM, *et al.* Airway nitric oxide levels in cystic fibrosis patients are related to a polymorphism in the neuronal nitric oxide synthase gene. Am J Respir Crit Care Med 2000; 162:2172-6.
[230] Grasemann H, Michler E, Wallot M, Ratjen F. Decreased concentration of exhaled nitric oxide (NO) in patients with cystic fibrosis. Pediatr Pulmonol 1997; 24:173-7.
[231] Accurso FJ, Sontag MK. Seeking modifier genes in cystic fibrosis. Am J Respir Crit Care Med 2003; 167:289-90.
[232] Gharib SA, Vaisar T, Aitken ML, Park DR, Heinecke JW, *et al.* Mapping the lung proteome in cystic fibrosis. J Proteome Res 2009; 8:3020-8.
[233] Stockley RA, Bayley DL, Unsal I, Dowson LJ. The effect of augmentation therapy on bronchial inflammation in alpha1-antitrypsin deficiency. Am J Respir Crit Care Med 2002; 165:1494-8.
[234] Camelier AA, Winter DH, Jardim JR, Barboza CE, Cukier A, *et al.* [Alpha-1 antitrypsin deficiency: diagnosis and treatment]. J Bras Pneumol 2008; 34:514-27.
[235] Stoller JK, Aboussouan LS. Alpha1-antitrypsin deficiency. Lancet 2005; 365:2225-36.
[236] Mahadeva R, Stewart S, Bilton D, Lomas DA. Alpha-1 antitrypsin deficiency alleles and severe cystic fibrosis lung disease. Thorax 1998; 53:1022-4.
[237] Jefferies H, Coster J, Khalil A, Bot J, McCauley RD, *et al.* Glutathione. ANZ J Surg 2003; 73:517-22.
[238] Gilliland FD, Gauderman WJ, Vora H, Rappaport E, Dubeau L. Effects of glutathione-S-transferase M1, T1, and P1 on childhood lung function growth. Am J Respir Crit Care Med 2002; 166:710-6.
[239] Baranova H, Perriot J, Albuisson E, Ivaschenko T, Baranov VS, *et al.* Peculiarities of the GSTM1 0/0 genotype in French heavy smokers with various types of chronic bronchitis. Hum Genet 1997; 99:822-6.

[240] Baranov VS, Ivaschenko T, Bakay B, Aseev M, Belotserkovskaya R*, et al.* Proportion of the GSTM1 0/0 genotype in some Slavic populations and its correlation with cystic fibrosis and some multifactorial diseases. Hum Genet 1996; 97:516-20.
[241] Kamada F, Mashimo Y, Inoue H, Shao C, Hirota T*, et al.* The GSTP1 gene is a susceptibility gene for childhood asthma and the GSTM1 gene is a modifier of the GSTP1 gene. Int Arch Allergy Immunol 2007; 144:275-86.
[242] Li YF, Gauderman WJ, Conti DV, Lin PC, Avol E*, et al.* Glutathione S-transferase P1, maternal smoking, and asthma in children: a haplotype-based analysis. Environ Health Perspect 2008; 116:409-15.
[243] Imboden M, Rochat T, Brutsche M, Schindler C, Downs SH*, et al.* Glutathione S-transferase genotype increases risk of progression from bronchial hyperresponsiveness to asthma in adults. Thorax 2008; 63:322-8.
[244] Aron Y, Polla BS, Bienvenu T, Dall'ava J, Dusser D*, et al.* HLA class II polymorphism in cystic fibrosis. A possible modifier of pulmonary phenotype. Am J Respir Crit Care Med 1999; 159:1464-8.
[245] Aron Y, Bienvenu T, Hubert D, Dusser D, Dall'Ava J*, et al.* HLA-DR polymorphism in allergic bronchopulmonary aspergillosis. J Allergy Clin Immunol 1999; 104:891-2.
[246] Eisen DP, Minchinton RM. Impact of mannose-binding lectin on susceptibility to infectious diseases. Clin Infect Dis 2003; 37:1496-505.
[247] Bouwman LH, Roep BO, Roos A. Mannose-binding lectin: clinical implications for infection, transplantation, and autoimmunity. Hum Immunol 2006; 67:247-56.
[248] Carlsson M, Sjoholm AG, Eriksson L, Thiel S, Jensenius JC*, et al.* Deficiency of the mannan-binding lectin pathway of complement and poor outcome in cystic fibrosis: bacterial colonization may be decisive for a relationship. Clin Exp Immunol 2005; 139:306-13.
[249] Garred P, Pressler T, Madsen HO, Frederiksen B, Svejgaard A*, et al.* Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis. J Clin Invest 1999; 104:431-7.
[250] Davies JC, Potter M, Bush A, Rosenthal M, Geddes DM*, et al.* Bone marrow stem cells do not repopulate the healthy upper respiratory tract. Pediatr Pulmonol 2002; 34:251-6.
[251] Muhlebach MS, MacDonald SL, Button B, Hubbard JJ, Turner ML*, et al.* Association between mannan-binding lectin and impaired lung function in cystic fibrosis may be age-dependent. Clin Exp Immunol 2006; 145:302-7.
[252] Dorfman R, Sandford A, Taylor C, Huang B, Frangolias D*, et al.* Complex two-gene modulation of lung disease severity in children with cystic fibrosis. J Clin Invest 2008; 118:1040-9.
[253] Taylor DR, Kennedy MA. Genetic variation of the beta(2)-adrenoceptor: its functional and clinical importance in bronchial asthma. Am J Pharmacogenomics 2001; 1:165-74.
[254] Ramsay CE, Hayden CM, Tiller KJ, Burton PR, Goldblatt J*, et al.* Polymorphisms in the beta2-adrenoreceptor gene are associated with decreased airway responsiveness. Clin Exp Allergy 1999; 29:1195-203.
[255] Turki J, Pak J, Green SA, Martin RJ, Liggett SB. Genetic polymorphisms of the beta 2-adrenergic receptor in nocturnal and nonnocturnal asthma. Evidence that Gly16 correlates with the nocturnal phenotype. J Clin Invest 1995; 95:1635-41.

[256] Large V, Hellstrom L, Reynisdottir S, Lonnqvist F, Eriksson P, *et al.* Human beta-2 adrenoceptor gene polymorphisms are highly frequent in obesity and associate with altered adipocyte beta-2 adrenoceptor function. J Clin Invest 1997; 100:3005-13.
[257] Buscher R, Eilmes KJ, Grasemann H, Torres B, Knauer N, *et al.* beta2 adrenoceptor gene polymorphisms in cystic fibrosis lung disease. Pharmacogenetics 2002; 12:347-53.
[258] Drumm ML, Konstan MW, Schluchter MD, Handler A, Pace R, *et al.* Genetic modifiers of lung disease in cystic fibrosis. N Engl J Med 2005; 353:1443-53.
[259] Hart MA, Konstan MW, Darrah RJ, Schluchter MD, Storfer-Isser A, *et al.* Beta 2 adrenergic receptor polymorphisms in cystic fibrosis. Pediatr Pulmonol 2005; 39:544-50.
[260] Steagall WK, Barrow BJ, Glasgow CG, Mendoza JW, Ehrmantraut M, *et al.* Beta-2-adrenergic receptor polymorphisms in cystic fibrosis. Pharmacogenet Genomics 2007; 17:425-30.
[261] Darrah R, McKone E, O'Connor C, Rodgers C, Genatossio A, *et al.* EDNRA variants associate with smooth muscle mRNA levels, cell proliferation rates, and cystic fibrosis pulmonary disease severity. Physiol Genomics 2010; 41:71-7.
[262] Stanke F, Becker T, Kumar V, Hedtfeld S, Becker C, *et al.* Genes that determine immunology and inflammation modify the basic defect of impaired ion conductance in cystic fibrosis epithelia. J Med Genet 2011.
[263] Trevisiol C, Boniotto M, Giglio L, Poli F, Morgutti M, *et al.* MBL2 polymorphisms screening in a regional Italian CF Center. J Cyst Fibros 2005; 4:189-91.
[264] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, *et al.* Initial sequencing and analysis of the human genome. Nature 2001; 409:860-921.
[265] Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, *et al.* Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev 1996; 13:19-50.
[266] Penque D. Proteomic biomarker discovery for the monogenic disease cystic fibrosis. Expert Rev Proteomics 2007; 4:199-209.
[267] Gruenert DC, Willems M, Cassiman JJ, Frizzell RA. Established cell lines used in cystic fibrosis research. J Cyst Fibros 2004; 3 Suppl 2:191-6.
[268] Pollard HB, Ji XD, Jozwik C, Jacobowitz DM. High abundance protein profiling of cystic fibrosis lung epithelial cells. Proteomics 2005; 5:2210-26.
[269] Pollard HB, Eidelman O, Jozwik C, Huang W, Srivastava M, *et al.* De novo biosynthetic profiling of high abundance proteins in cystic fibrosis lung epithelial cells. Mol Cell Proteomics 2006; 5:1628-37.
[270] Davezac N, Tondelier D, Lipecka J, Fanen P, Demaugre F, *et al.* Global proteomic approach unmasks involvement of keratins 8 and 18 in the delivery of cystic fibrosis transmembrane conductance regulator (CFTR)/deltaF508-CFTR to the plasma membrane. Proteomics 2004; 4:3833-44.
[271] Brouillard F, Bensalem N, Hinzpeter A, Tondelier D, Trudel S, *et al.* Blue native/SDS-PAGE analysis reveals reduced expression of the mClCA3 protein in cystic fibrosis knock-out mice. Mol Cell Proteomics 2005; 4:1762-75.
[272] Bensalem N, Ventura AP, Vallee B, Lipecka J, Tondelier D, *et al.* Down-regulation of the anti-inflammatory protein annexin A1 in cystic fibrosis knock-out mice and patients. Mol Cell Proteomics 2005; 4:1591-601.

[273] Carvalho-Oliveira IM, Charro N, Aarbiou J, Buijs-Offerman RM, Wilke M*, et al.* Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009; 8:3606-16.
[274] Sloane AJ, Lindner RA, Prasad SS, Sebastian LT, Pedersen SK*, et al.* Proteomic analysis of sputum from adults and children with cystic fibrosis and from control subjects. Am J Respir Crit Care Med 2005; 172:1416-26.
[275] Pedersen SK, Sloane AJ, Prasad SS, Sebastian LT, Lindner RA*, et al.* An immunoproteomic approach for identification of clinical biomarkers for monitoring disease: application to cystic fibrosis. Mol Cell Proteomics 2005; 4:1052-60.
[276] Gray RD, Macgregor G, Noble D, Imrie M, Dewar M*, et al.* Sputum Proteomics in Inflammatory and Suppurative Respiratory Diseases. Am J Respir Crit Care Med 2008.
[277] Donaldson SH, Bennett WD, Zeman KL, Knowles MR, Tarran R*, et al.* Mucus clearance and lung function in cystic fibrosis with hypertonic saline. N Engl J Med 2006; 354:241-50.
[278] MacGregor G, Gray RD, Hilliard TN, Imrie M, Boyd AC*, et al.* Biomarkers for cystic fibrosis lung disease: application of SELDI-TOF mass spectrometry to BAL fluid. J Cyst Fibros 2008; 7:352-8.
[279] von Bredow C, Birrer P, Griese M. Surfactant protein A and other bronchoalveolar lavage fluid proteins are altered in cystic fibrosis. Eur Respir J 2001; 17:716-22.
[280] Blouquit S, Regnier A, Dannhoffer L, Fermanian C, Naline E*, et al.* Ion and fluid transport properties of small airways in cystic fibrosis. Am J Respir Crit Care Med 2006; 174:299-305.
[281] Roxo-Rosa M, da Costa G, Luider TM, Scholte BJ, Coelho AV*, et al.* Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006; 6:2314-25.
[282] Gomes-Alves P, Imrie M, Gray RD, Nogueira P, Ciordia S*, et al.* SELDI-TOF biomarker signatures for cystic fibrosis, asthma and chronic obstructive pulmonary disease. Clin Biochem 2010; 43:168-77.
[283] Wright JM, Merlo CA, Reynolds JB, Zeitlin PL, Garcia JG*, et al.* Respiratory epithelial gene expression in patients with mild and severe cystic fibrosis lung disease. Am J Respir Cell Mol Biol 2006; 35:327-36.
[284] Tambor V, Fucikova A, Lenco J, Kacerovsky M, Rehacek V*, et al.* Application of proteomics in biomarker discovery: a primer for the clinician. Physiol Res 2010; 59:471-97.
[285] Apweiler R, Aslanidis C, Deufel T, Gerstner A, Hansen J*, et al.* Approaching clinical proteomics: current state and future fields of application in cellular proteomics. Cytometry A 2009; 75:816-32.
[286] Kiechle FL, Zhang X. The postgenomic era: implications for the clinical laboratory. Arch Pathol Lab Med 2002; 126:255-62.

Chapter II

Serum Proteomics Signature of Cystic Fibrosis Patients: a Complementary 2-DE and LC-MS/MS Approach

Manuscript published in Journal of Proteomics with minor alterations

Charro N, Hood BL, Faria D, Pacheco P, Azevedo P, Lopes C, de Almeida AB, Couto FM, Conrads TP, Penque D, Journal of Proteomics 2011, **74**(1); 110-26.

ABSTRACT

Complementary 2D-PAGE and 'shotgun' LC-MS/MS approaches were combined to identify medium and low-abundant proteins in sera of Cystic Fibrosis (CF) patients (mild or severe pulmonary disease) in comparison with healthy CF-carrier and non-CF carrier individuals aiming to gain deeper insights into the pathogenesis of this multifactorial genetic disease.

78 differentially expressed spots were identified from 2D-PAGE proteome profiling yielding 28 identifications and postulating the existence of post-translation modifications (PTM).

The 'shotgun' approach highlighted altered levels of proteins actively involved in CF: abnormal tissue/airway remodeling, protease/antiprotease imbalance, innate immune dysfunction, chronic inflammation, nutritional imbalance and *P. aeruginosa* colonization. Members of the apolipoproteins family (VDBP, ApoAI, ApoB) presented gradually lower expression from non-CF to CF carrier individuals and from those to CF patients, results validated by an independent assay. The multifunctional enzyme NDKB was identified only in the CF group and independently validated by WB. Its functions account for ion sensor in epithelial cells, pancreatic secretion, neutrophil-mediated inflammation and energy production, highlighting its physiological significance in the context of CF.

Complementary proteomics-based approaches are reliable tools to reveal pathways and circulating proteins actively involved in a heterogeneous disease such as CF.

Keywords:

Cystic Fibrosis, Serum Proteome Profiling, 2DE-MALDI-TOF/TOF MS, Shotgun LC-MS/MS, Label Free Proteomics

1. BACKGROUND

Cystic Fibrosis (CF) is the most common autosomal recessive monogenic disease in Caucasians caused by mutations in the gene coding for the CF-transmembrane conductance regulator (CFTR) protein [1-3]. CF is classically characterized by progressive bronchiectatic lung disease, pancreatic exocrine insufficiency and male infertility. Although a broad spectrum of disease phenotypes exists [4], airway inflammation and recurrent infections leading to respiratory failure are the major causes of morbidity and mortality among CF patients and, therefore, the main focus of clinical research [2,3,5]. The impairment of CFTR alone does not completely explain the complex spectrum of CF clinical phenotypes; hence characterization of proteins with new and undefined roles in CF pathophysiology is critical for more accurate diagnosis and/or prognosis of this disease [6] as there is accumulating evidence that genetic factors other than CFTR significantly influence severity of CF-related lung disease [7].

Several studies aim to gain a better understanding of proteins and their interactions within the lung microenvironment using sputum [8,9] and bronchoalveolar lavage fluid (BALF) [10-13] while identifying proteins to serve as biomarkers for diagnosis/prognosis of CF. Ultimately, the candidate biomarker(s) will be translated into clinical assays for rapid, non-invasive and inexpensive analysis of blood, saliva or urine and serum appears as one of the ideal sources for biomarkers due to its low invasiveness, minimal cost, easy sample collection and processing [14] and interaction with the majority of the body tissues, thus virtually containing any potentially detectable combination of all the differentiated sub-proteomes of the body [15-18]. However, the complexity and wide protein concentration range, which spans over more than 10 orders of magnitude, makes serum one of the most challenging specimens to analyze as a whole [17,19]. It is well known that biomarkers are low-abundant proteins with concentrations in serum/plasma in the ng/mL to pg/mL range [20] which comprise less than 1% of the total plasma proteome [19,21]. To detect these lower abundance proteins in serum, efficient throughput depletion strategies and post depletion fractionation are necessary [16]. Among the various depletion approaches that have been described [21], the immunodepletion-based methods provide the highest depletion selectivity and reproducibility [22] that avoid the masking effect of the high abundant proteins and allow for more comprehensive identification/quantification of the low abundant ones.

The first attempt to discriminate between sera of CF patients and non-CF healthy individuals utilized a microarray-based clinical test of 507 antibodies against low abundant intracellular proteins [23]. A CF-specific serum proteomics signature was identified and comprised proteins mediators of inflammation from the NFκB signalling pathway. In another investigation, an immunoproteomic approach [24] was utilized and resulted in the identification of antibodies in serum of CF patients against inflammatory and defense proteins.

Protein identification has become routine with modern mass spectrometers and accurate and reliable quantification of differentially expressed proteins is essential to validate the candidate biomarkers. Wang *et al*. and others [25,26] reported a quantification method that relies on direct comparison of peptide peak areas between LC-MS runs without any isotopic labelling. This called label-free approach showed that the intensities of peptide peak signals correspond nearly linearly to their concentrations in the sample and that the ratios of peak areas between different LC-MS runs can be used for quantification as they reflect relative quantities among samples [15,25].

Here we have utilized a complementary analysis of serum proteins by 2D-PAGE and a "shotgun" approach using LC-MS/MS to identify differentially abundant proteins in sera from patients with CF. The analysis of the 2D-PAGE serum proteome profiles resulted in the recognition of 78 differentially expressed spots from which 28 proteins were identified and therefore postulating the presence of PTMs. The shotgun proteomics approach resulted in the confident identification of over 400 serum proteins from the groups under study with (73±5) differentially expressed proteins when comparisons of any 2 groups were performed. The combined use of these proteomic approaches enabled a complementary view into CF pathogenesis highlighting processes and proteins relevant in CF lung disease with potential clinical interest.

2. MATERIAL AND METHODS

2.1 Patients characterization and sample collection

Ethical approval was obtained from Instituto Nacional de Saúde Dr. Ricardo Jorge and Hospital de Santa Maria.

Thirty five CF patients with heterogeneous mutations in the CFTR gene under clinical surveillance in Pulmonology's Division were enrolled for this study. All CF

patients were free of acute phase of inflammation and under standard respiratory treatments, namely nebulized recombinant human dornase alpha (rhDNAse), tobramycin and azithromycin to reduce sputum viscoelasticity, control respiratory infection and exacerbation and inhibit bacterial proliferation, respectively.

After informed consent, peripheral blood was collected by veinipuncture into anticoagulant-free tube and allowed to clot for 30 min at ambient temperature. Serum was isolated by centrifugation at 2000 x *g* for 10 min and immediately aliquoted and stored at -80 ºC. To emphasize proteomic differences among the groups studied while eliminating potential individual contributions, equal volumes of serum from all CF patients enrolled or those reflecting different stages of CF lung disease (with mild or severe respiratory disease) classified based on forced expiratory volume in 1 second (FEV_1) were pooled according to [27] and summarized in Table II.1.

By analyzing patients with a multiplicity of mutation's types, we aim to praise significant differences between CF and non-CF groups. Pooled serum samples from age- and gender-matched healthy individuals (n=31) and carriers of one mutated *CFTR* allele (n=21) were used for comparison purposes. All healthy individuals were previously genotyped for the most common mutations in the Portuguese population [28] to confirm that they were non-CF carriers and clinically evaluated (questionnaires, chest auscultation and spirometry) to address respiratory function ensuring a $FEV_1 \geq$ 80% of vital capacity.

Table II.1: Cystic Fibrosis (CF) patients and Carriers demographics and stratification into disease severity based on clinical criteria and pulmonary function accessed by spirometry [29]. Healthy individuals were chosen to match similar demographic conditions as the CF patients.

	Group				
	Controls	CF Patients			Carriers
		Total	Mild	Severe	
n	31	35	11	24	21
Age (mean ± SD)	(30±8) y	(27±8) y	(27±9) y	(27±8) y	(45±11) y
Gender	20 F 11 M	20 F 15 M	7 F 4 M	13 F 11 M	13 F 8 M
Genotype ΔF/ΔF (%)	NA	7 (20%)	0 (0%)	7 (20%)	NA
Genotype ΔF (%)	NA	22 (63%)	8 (73%)	14 (58%)	10 (48%)
FVC % (mean ± SD)	93 ± 11	77 ± 24	89 ± 23	71 ± 23	---
FEV_1 % (mean ± SD)	91 ± 12	60 ± 27	80 ± 24	51 ± 24	---
Colonization P. aeruginosa (%)	---	60	45	67	---
BMI (mean ± SD)	---	21.1 ± 3.5	22.9 ± 3.6	20.3 ± 3.2	---

2.2 Depletion of the 6 most abundant proteins in serum by an immunoaffinity-based method

There are a number of basic techniques that have been used to eliminate the high abundance proteins in serum. Here, we used the Multiple Affinity Removal System (MARS) spin cartridge (P/N 5188-5230, 0.45mL resin bed, Agilent Technologies, Palo Alto, CA) with binding capacity for 16μL of crude human serum to remove the six most abundant proteins in serum (albumin, IgA, IgG, α-1-antitrypsin, transferrin and haptoglobin). According to manufacturer's instructions, pooled sera were diluted 8-fold with buffer A (P/N 5185-5987, Agilent Technologies) supplemented with Protease Inhibitor Cocktail Mix (Sigma-Aldrich) diluted 1000-fold (v/v) and filtered through a 0.22μm pore size filter by spinning at 2320 x *g* for 5 min at 4 ºC for removal of particles. Approximately 200 μL of the diluted and filtered serum were applied on the MARS cartridge and spun at 100 x *g* for 1.5 min at 4 ºC. Column was washed 2 times with buffer A followed by spinning at 100 x *g* for 2 min at 4 ºC and all the 3 flow-through fractions were combined and kept on ice all the time. Bound proteins were eluted with 2mL of buffer B (P/N 5185-5988, Agilent Technologies) and the cartridge

was re-equilibrated with 4 mL of buffer A, which was then ready for a new cycle of immunodepletion as described.

2.3 Desalting and concentration of the flow-through fractions by centrifugal ultrafiltration

Flow-through fractions of the same sample from several cycles of immunodepletion were pooled, buffer-exchanged and concentrated using 4 mL spin concentrators with 5 kDa molecular weight cut off (Amicon Ultra 4, MWCO 5000Da, Millipore). After membrane activation according to manufacturer's protocol, the sample was filtered at 2000 x *g* at 4 ºC until ≈ 80% of the initial volume and gradually buffer-exchanged into 25 mM ammonium bicarbonate (NH_4HCO_3) pH 8.4 by 3 rounds of addition of the buffer (ensuring removal of ≈ 99% of salts), with centrifugation at 2000 x *g* for 30 min at 4 ºC until a final volume of ≈ 300 µL. Removal of the high abundant proteins was confirmed by the analysis of the flow-through and bound fractions by 1-D SDS-PAGE *(data not show)*. Protein concentrations were established using a BCA™ protein assay kit (Pierce). These samples were stored at -80 ºC until further use for 2-DE and LC-MS/MS analysis.

2.4.1 2-D Electrophoresis

2-DE was carried out essentially as previously described [5]. Approximately 600 µg of protein from the immunodepleted serum samples were resuspended in IEF buffer [7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 60 mM (w/v) DTE, 0.75 % (w/v) ampholine 3.5-10.0, 0.25 % (w/v) ampholine 4.0-6.0) before loading onto precast 18 cm-long Immobiline™ DryStrips (GE HealthCare) with a non-linear wide range pH gradient (pH 3-10). Proteins were focused for a total of ≈ 100 kVh, during which the voltage was gradually increased up to 8000 V for a total of 39 h.

The second dimension, SDS-PAGE, was resolved overnight on 8-16 % linear gradient polyacrylamide gels at 10 ºC with constant 3.5 W *per* gel in an Ettan DALTwelve System (GE HealthCare). After electrophoresis, gels were stained with Coomassie Brilliant Blue (CBB) as previously described [30]. A total of four 2DE gels *per* group were run.

2.4.2 2-DE maps analysis

In total, 20 CBB-stained 2-DE gels were digitalized on ImageScanner (GE HealthCare) and the obtained images converted into greyscale 14-bit TIFF files, which were then analyzed on Progenesis SameSpotst^{TM} version 2.0 (NonLinear Dynamics) image analysis program. Normalized spot volumes, *i.e.*, the volume of each spot over the volume of all spots in the gel, were used for comparison of the different clinical groups. Quantitative differences were assessed using an analysis of variance (ANOVA) test for *n* observations, where *n* is the number of groups under analysis. Spots displaying at least a 1.5 fold difference in protein expression were considered statistically significant for $p<0.05$ and excised for identification by MS.

2.5 In-gel digestion and protein identification by combined MS + MS/MS

Differentially expressed spots were excised from the gel and submitted to in-gel digestion prior to identification by MS as described [31,32]. Tryptic peptides were deposited on a 192-well MALDI plate with 5 mg/mL α-CHCA (LaserBiolabs, 1:1) in 0.1 % TFA/60 % ACN (v/v) and allowed to co-crystallize at ambient temperature.

Peptides were analyzed by MALDI-TOF/TOF MS (ABI 4700 Proteomics Analyzer, Applied Biosystems). Internal calibration, data acquisition, processing and interpretation were carried out according to [33] using vendor's software packages (4000 Series Explorer software v3.0 RC1, ABI and GPS Explorer software v3.5, ABI) and a local copy of Mascot search engine (Version 2.0).

For combined MS+MS/MS analysis, all peptide mass values are considered monoisotopic, a MS mass tolerance was set at 50 ppm and a MS/MS fragment tolerance set at 0.25 Da. Trypsin was assigned as the digestion enzyme, a single missed cleavage site was allowed and carboamidomethylation of cysteinyl and oxidation of methionyl residues were assumed as fixed and variable modifications, respectively. A taxonomic restriction to human protein sequences was included. For MS, all peaks with a signal to noise ratio (S/N) greater than 5 and for MS/MS all peaks with a S/N greater than 3 were included in the database search against the SwissProt database (Swissprot 2008.06.09, 385721 sequences, 19630 human sequences). The criteria used to accept the identification was significant homology scores achieved in Mascot ($p< 0.05$).

2.6 Analysis of the flow-through fractions by nanoRPLC-MS/MS

200 μg of immunodepleted serum proteins from the several groups under study were digested with sequence-grade modified trypsin (Promega, Madison, WI, USA) in 25 mM NH_4HCO_3 at an enzyme to protein ratio of 1:50 and each sample was incubated for 12-16 h at 37 ºC. Digestion was stopped by adding TFA to a final concentration of 0.1 % (v/v) and samples were desalted by solid phase extraction (High Performance Extraction Disk Cartridges, Empore, 3M) according to the manufacturer's protocol. Digests were filtered 3 times to maximize recovery and the peptides were eluted twice with 0.1 % TFA/60 % ACN (v/v) before vacuum centrifugation to dryness.

Peptides were dissolved in 0.1 % TFA (v/v) and analyzed in triplicate by nanoflow RPLC coupled online with a LIT mass spectrometer (LTQ-XL, ThermoFisher Scientific Inc., San Jose, CA). Separation of the peptide digests was performed using integrated ESI-capillary RP columns (75 μm ID x 360 μm OD x 100 mm length) packed in-house with 5 μm 300 Å pore size C18 RP stationary phase (Jupiter, Phenomenex, Torrance, CA). Mobile phase flow was supplied by a nanoflow HPLC system (Ultimate 3000, Dionex Corporation, Sunnyvale, CA). After injection, the column was washed with 98 % mobile phase A [0.1 % FA (v/v)] for 30 min and peptides were eluted by development of a linear gradient from 2 % to 42 % mobile phase B [0.1 % FA (v/v) in ACN] in 140 min, then to 98 % B in an additional 20 min, all at a constant flow rate of 250 nL/min. The MS was operated in data-dependent MS/MS mode in which each full MS scan (scan range of *m/z* 350–1800) was followed by seven MS/MS scans where the seven most abundant peptide molecular ions dynamically determined from the MS scan were selected for tandem MS using a relative CID energy of 30 %. Dynamic exclusion was utilized to minimize redundant selection of peptides for CID.

2.7 Data Analysis

The online tool FindMod (ww.expasy.org/tools/findmod.html) was used to predict post translational modifications (PTMs) on unassigned peaks from MALDI-TOF/TOF MS analysis. This tool calculates the differences between experimentally determined peptide masses and theoretical masses calculated from a specific protein sequence and allocates them to the 22 types of PTMs existent in the UniProt knowledge base.

Tandem mass spectra were searched against the UniProt human proteome database (http://www.expasy.org; 01/08 release) using SEQUEST™ (ThermoFisher Scientific Inc.). Peptides were considered legitimately identified if they achieved specific charge state and proteolytic cleavage-dependent cross-correlation (XC) scores of 1.9 for $[M+H]^{1+}$, 2.2 for $[M+2H]^{2+}$, and 3.5 for $[M+3H]^{3+}$, and a minimum delta correlation score (ΔCn) of 0.08. Protein abundances were calculated by comparing total spectral counting values across analysed conditions.

PIKE (Protein Information Knowledge Extractor, http://proteo.cnb.uam.es:8080/pike/) software was used to combine the available information about each protein using different databases. This open-source tool retrieves a set of biological information such as gene name, subcellular location, tissue specificity, function and association with disease.

Ingenuity Pathways Analysis (Ingenuity Systems®, www.ingenuity.com, IPA) software was used to integrate the identified proteins into signaling pathways with biological meaning. Datasets containing gene identifiers and corresponding expression values were uploaded and mapped to their gene object in the IPA's knowledge base. Networks of these genes were then algorithmically generated based on their connectivity and interactions. Functional analysis identified the most significant molecular and cellular functions and/or disorders to each dataset.

Reactome [34] (www.reactome.org), a freely available curated pathway database, was also used for integrate identified proteins into networks with biological significance.

In order to characterize each dataset, a list of all GO terms (separated by GO type) annotated to proteins in that dataset was obtained from the GOA-UniProt database (as of the December 2007 release) and the frequency of annotation of each term within the dataset was calculated. The information content (*IC*) of each GO term was calculated to measure its specificity (*i.e.* the level of detail with which it characterizes the set), according to the expression:

$$IC(t) = \frac{-\log(f_t)}{\log(N_{Total})} \text{ , (\textbf{Eq. 1})}$$

where f_t is the frequency of annotation of the term within the full GOA-UniProt database (an estimation of its probability of annotation) and N_{Total} is the total number of annotations of the respective GO type in the GOA-UniProt database (a normalizing factor, so that the *IC* is obtained in the 0-1 scale).

To measure the statistical significance of the occurrence of each term within each dataset, an *e*-value was calculated by estimating the probability of a random set of proteins having for that term a frequency of annotation greater than the frequency obtained in the real set.

Assuming that the proteins in a random dataset are independent, the probability of a given number of proteins having a specific term (successes) and the remaining proteins not having it (failures) is given by a binomial distribution:

$$P[X = x \mid f_t] = \binom{n}{x} \times f_t^{\,x} \times (1 - f_t)^{(1-x)} \text{ , (\textbf{Eq. 2})}$$

where x is the number of successes (*i.e.* the number of proteins in the dataset annotated with the term), n is the size of the dataset, and f_t is an estimation of the probability of annotation of the term, as discussed above.

Thus, the *e*-value is given by the expression:

$$P[X > x \mid f_t] = \sum_{i=x+1}^{i=n} \binom{n}{i} \times f_t^{\,i} \times (1 - f_t)^{(1-i)} \text{ , (\textbf{Eq. 3})}$$

(Note: The *e*-values were calculated using Microsoft Excel's function for the binomial distribution).

With these parameters, the characterization of each dataset is complete, allowing us to identify the GO terms which are significantly overrepresented in the dataset and the most informative among those.

2.8 Biochemical validation

30 µg of serum's depleted fractions of the groups under analysis were separated on 4-12 % (w/v) polyacrylamide gels, transferred to nitrocellulose membranes (Schleicher&Schuell) and analyzed by western blot, each triplicate analyzed in independent gels. Membranes were probed with mouse monoclonal anti-NME2

(Abcam) diluted 1:1000 in PBS supplemented with 5 % (w/v) fat free milk for 2 h at ambient temperature and developed using chemiluminescence-ECL (Pierce). The abundance of selected protein was calculated from densitometry of immunoblots using Progenesis PG200v2006.

Alpha-1-antitrypsin, complement C4 and ceruloplasmin were measured in individual sera by a nephelometric assay using Beckman Coulter Array® Systems kits for the specified proteins and Immage® 800 instrument (Beckman Coulter). Apolipoproteins A-I and B were determined in individual sera from all the individuals enrolled by an immunoturbidimetric assay using Cobas® Integra Tina-quant kit and Cobas® Integra 400 instrument (Roche). Statistical analysis between groups was performed using the ANOVA test or Kruskal-Wallis test and values of $p<0.05$ were assumed as significant. Tukey´s Test or Dunn's Multiple Comparison Test were performed across all groups to investigate differences between them using GraphPad Prism (v.4.02 GraphPad software).

3. RESULTS

Serum plays a central role in biomarker discovery studies because it is likely to contain valuable pathophysiological information as it perfuses all of the body's tissues [15,16,18,35]. However, due to wide dynamic range of protein concentration, prefractionation and enrichment techniques are generally applied to enable detection of proteins at lower abundance levels. Here, we employed complementary gel- and solution-based MS approaches to perform a comparative proteomics analysis of immunodepleted serum from CF patients in comparison with healthy non-CF individuals and CF carriers. Enrolled individuals were grouped according to their genetic background (mutational-based analysis) and assigned as controls (*e.g.*, no mutations found in the *CFTR* gene and normal pulmonary function), carriers of one mutated *CFTR* allele and CF patients bearing two identified mutations. The CF patients were further sub-grouped as mild or severe CF according to their pulmonary function assessed by spirometry and clinical criteria [27]. Individual serum samples were pooled prior to depletion to emphasize proteomics differences while diluting potential individual contributions (Table II.1).

3.1 2-DE of depleted serum fraction

The immunodepleted protein fractions were resolved by 2-DE using IPG strips with a wide pH range (3-10 NL) in the 1st dimension and 8-16% polyacrylamide gradient gels in the 2nd dimension and the obtained maps were visualized by CBB staining (Figures II.1.1 and II.1.2). Four gel replicas *per* group under comparison were run and gel images were analyzed using Progenesis SameSpots™ software. After spots alignment across gels, statistical analysis was performed on the normalized volume differences. Comparison of 2D-gel images of controls with carriers and CF patients resulted in recognition of 25 and 10 differentially abundant protein spots (fold ± 1.5; p(ANOVA) <0.05), respectively. 19 differentially abundant spots were found between carriers and CF patients (Table II.1.1SD). We also considered spots with a fold difference of ± 1.3 for excision and identification due to spots proximity or local low resolution. Spots were excised, digested and 23 unique proteins were identified by MALDI-TOF/TOF MS (Figure II.1.1 and Table II.2.1SD). When CF patients with different stages of lung disease were compared, 24 differentially abundant protein spots were found (Table II.1.2SD), corresponding to the identification of 16 unique proteins, most of them also identified when controls, CF carriers and patients were compared (Figure II.1.2 and Table II.2.2SD).

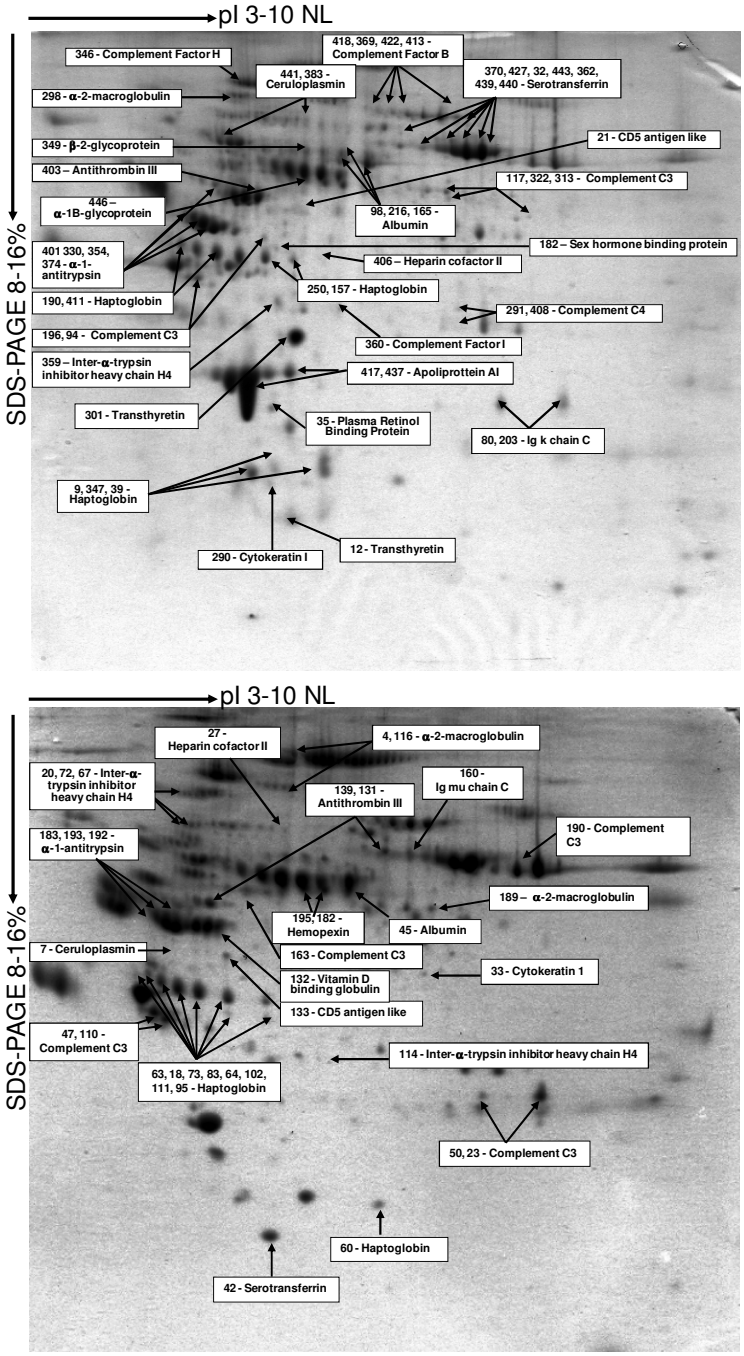


Figure II.1.1 and II.1.2: 2-DE reference map of serum depleted from the six most abundant proteins from the mutational-based analysis and respiratory-based analysis, respectively, with the indication of the differentially expressed protein (ANOVA test, $p<0.05$; 4 gel replicas per group, total 20 gels). Highlighted protein spots were identified by MS (Tables II.2.1 and II.2.2SD).

Identified proteins such as complement C4, complement factor B and I, apolipoprotein A-I, ceruloplasmin or transthyretin belong to the category of medium abundant proteins whereas CD5 antigen-like, inter-α-trypsin inhibitor heavy chain H4, hemopexin or β-2-glycoprotein are believed to belong to the class of borderline or low

abundant proteins [36], showing that depletion is a suitable sample preparation to go deeper into the proteome of such complex sample. Different spots resulted in the identification of the same protein and therefore we postulate the existence of PTMs or protein isoforms. Retrieved results from FindMod showed high probability of some proteins being modified (Table II.3.1SD). α-2-macroglobulin, for example, is predicted to be modified in different stages of CF lung disease but unmodified when comparing controls with carriers and CF, suggesting that PTMs pattern might be significant in the severity of the pathology. The putative deamidated form (spot 4) was found to be up regulated in CF patients with severe lung disease while one of the putative phosphorylated specie (spot 116) was down regulated in the same condition. Complement C3 is predicted to have a trimethylated modified lysine residue (spot 94) that is down regulated in carriers in comparison with controls and CF patients (similar levels in these 2 groups) and an acetylated form (spot 196) down regulated in carriers and CF patients in comparison with healthy individuals (similar levels in carriers and CF patients). No PTM was predicted for this protein when different stages of CF lung disease status were considered. Haptoglobin was by far the protein identified in a bigger number of spots, 15 in total. Predicted PTMs for this protein include deamidation, mono- or dimethylation and modification of lysine residues. Other examples of proteins potentially modified include serotransferrin, albumin, inter-α-trypsin inhibitor heavy chain H4, complement C4, complement factor I and β-2-glycoprotein. Differences in the patterns of modifications might be responsible for different functions and/or activities of proteins with virtual relevance in the disease phenotype.

3.2 1D-RPLC-MS/MS

Immunodepleted serum samples were digested with trypsin and peptides analyzed recursively by LC-MS/MS. Protein abundances were inferred by the label-free methodology of spectral counting [25,26]. The identified proteins in the control group were used for both mutational and respiratory-based comparisons. The total spectral counts (TSC) was very similar across samples (median = 1759; relative standard deviation = 15%), reinforcing the ability to reliably compare protein abundances. The false discovery rate (FDR) calculated by searching the data against a decoy database was 5.3%.

A total of 1593 and 1628 proteins were identified in mutational and respiratory-based analysis, respectively (Tables II.4.2SD, II.4.3SD and II.5SD). However, only 327 proteins were identified by at least 2 peptides and considered for relative quantification in both types of analysis [20% of total identifications]. Of those, 72 and 79 were found differentially expressed when carriers and CF patients were compared with healthy individuals, respectively, while 73 were differentially expressed in CF compared with carriers. 68 proteins presented altered abundances when CF patients with severe or mild pulmonary disease were compared (Table II.2).

Table II.2: Total proteins identified in the depleted fraction of serum by 1D RPLC-MS/MS using a label-free approach. The number of differentially expressed proteins in mutational-based analysis was obtained in comparison with Controls while in respiratory-based analysis only the two stages of pulmonary function were considered.

Groups	**# identified proteins**	**≥ 2 peptides (%)**	
Mutational Analysis	1593	327 (21%)	**≠ Expressed**
- Controls	763	187 (25%)	---
- Carriers	746	175 (23%)	72 (↑ 32 and ↓ 40)
- CF Patients	569	151 (27%)	79 (↑ 34 and ↓ 45) (with Controls) 73 (↑ 27 and ↓ 46) (with Carriers)
Respiratory Function	1628	327 (20%)	**≠ Expressed**
- Mild CF	664	165 (25%)	---
- Severe CF	722	163 (23%)	68 (↑ 25 and ↓ 43)

Ten proteins were simultaneously identified by 2-DE and LC-MS/MS (Figure II.2).

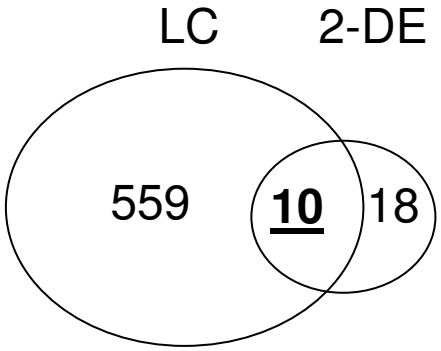


Figure II.2: Total proteins identified in LC-MS/MS (by at least 2 peptides) and in 2-DE experiments.

3.3 Subcellular Location and Functional Characterization

The obtained datasets were submitted to PIKE for cellular location. Within the annotated proteins, their cellular location is found to be spread throughout several cellular compartments, with cytoplasmic, nuclear and membrane proteins similarly represented. A significant amount of proteins had no information annotated (37%) whereas 5% were classified as mitochondrial or as belonging to ER or Golgi (Figure II.3), compartments typically associated with proteins synthesis and/or maturation.

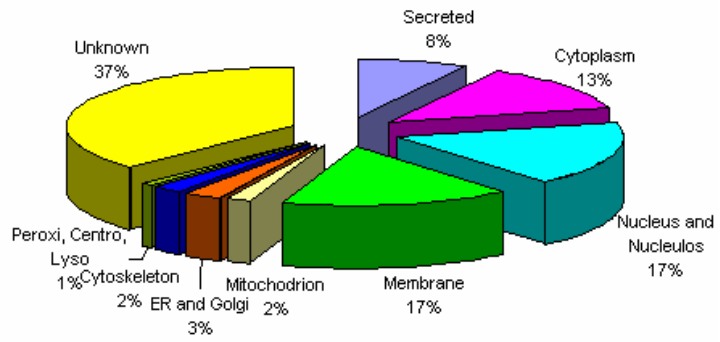


Figure II.3: Graphical representation of the subcellular location of the identified proteins by label-free LC-MS/MS according to PIKE software.

Functional semantic similarity between proteins (GO terms) was calculated by ProteInOn [37]. A statistical score was applied to quantify the representativeness of a GO term in the set of proteins and rank the results according to how well a given GO term describes that set. This tool highlighted important biological processes directly related with CF (Table II.6SD), namely ER overload response, cellular response to stress, exacerbated inflammatory response or wound healing, functions highly associated with CF condition.

Since our aim was to compare three conditions at a time (controls in comparison with carriers and those with CF patients or controls with CF patients with mild or severe lung disease), graphical expression profiles were establish to provide a wider overview of the expression tendencies among compared groups (Tables II.7.1SD and II.7.2SD). Only 2 proteins showed an increased expression from controls to carriers and CF patients. Vinculin is increased in relative abundance by 2- and 20-fold in carriers and CF, respectively, and might play a role in cell morphology, locomotion and cell adhesion. Aspartyl-tRNA synthetase is a phosphoprotein involved in protein

biosynthesis and in protein complex assembly and is increased in relative abundance by 1.5- and 2-fold in carriers and CF patients, respectively.

As respiratory function is concerned, proteins involved in antioxidant defense (selenoprotein P and serum paraoxonase/arylesterase 1) and in lung development (molecular chaperone TCP1 and β-catenin) were found to have lower abundances in patients with severe lung disease.

Several proteins involved in lipid metabolism or related with nutritional status were down regulated in carriers and CF patients, such as apolipoproteins or vitamin D-binding protein while variable expression profiles of the several complement components were observed. Anti-inflammatory proteins (inter-α-trypsin inhibitors or endothelin-1 receptor) show less abundance in CF patients' serum although with similar levels when lung disease severity was considered.

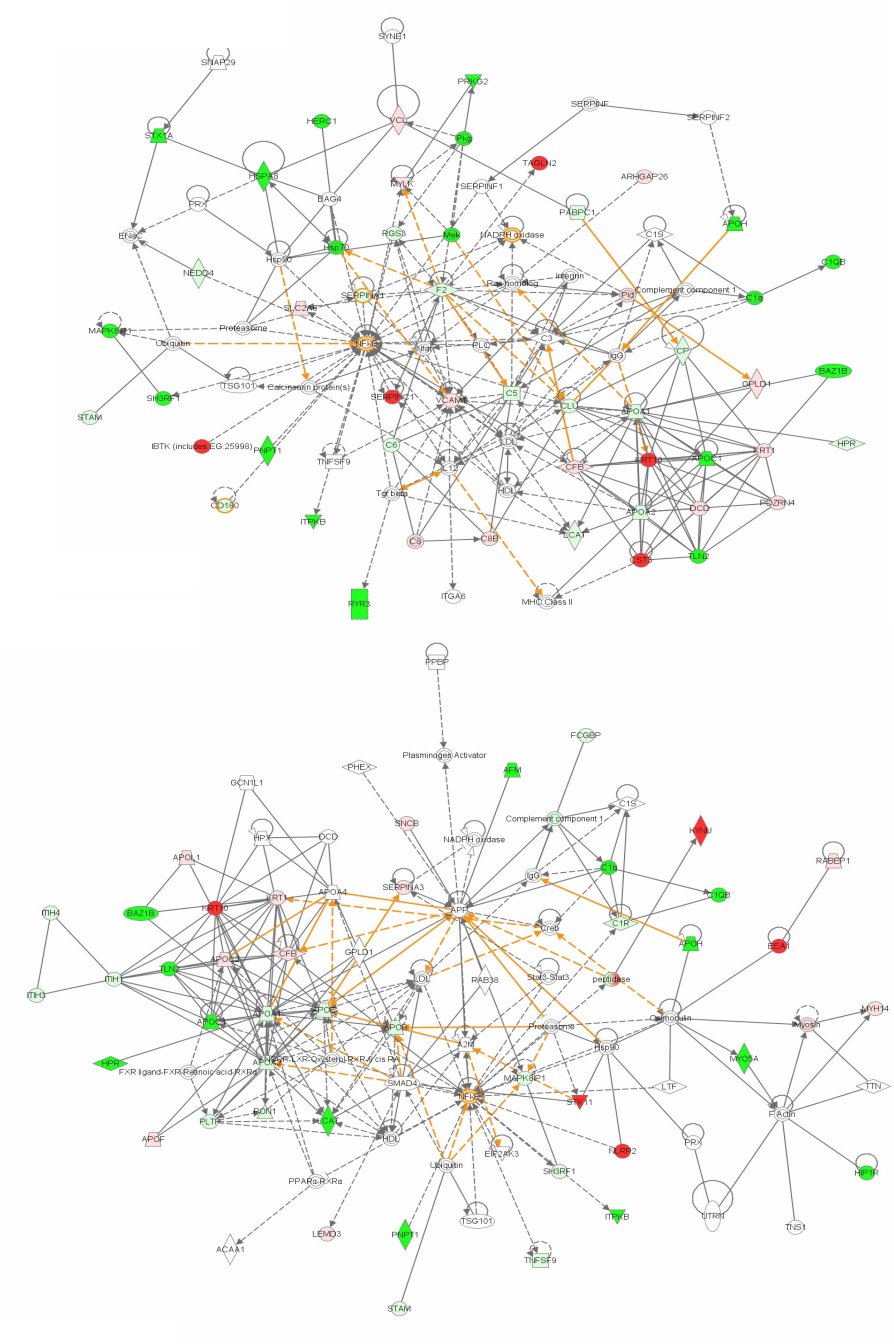
3.4 Signaling pathways and functional analysis

Ingenuity Pathway Analysis is highly-curated and comprehensive software used for integration of proteins into networks and pathways with biological meaning. Among those pathways generated, the 10 most significant ones were further explored (Figure II.1.1SD and II.1.2SD) and identified proteins with corresponding expression tendencies are highlighted. Processes such as antigen presentation, cell and humoral immune response, cell-to-cell signaling and interaction and organ morphology were identified as significant to the biological question under study. Confirming these results are the biological pathways retrieved by Reactome which consisted mainly in activation of complement cascade, hemostasis and metabolism of lipids and proteins.

Several proteins identified here were already described as NFκB interactors, an important signaling pathway that strongly contributes to the pro-inflammatory phenotype in CF. Although NFκB was not identified in our study, several proteins known to interact with it were identified, such as CD180 antigen, BAG family molecular chaperone regulator 4 and tumor necrosis factor ligand superfamily member 9 (Figures II.4.1 and II.4.2).

Members of the complement system and the apolipoprotein's family were also identified and described as modulating this pathway. Proteins belonging to HSP70 and HSP90 systems or ubiquitin were also associated with these relevant biological

processes, indicating a propitious environment related to protein maturation and/or degradation. Proteins generally associated with inflammatory processes were identified, such as inter-α-trypsin inhibitor heavy chain H1, H3 and H4, ceruloplasmin or the complement system, all of which may be important because in CF disease pathogenesis (Figure II.1.1SD and II.1.2SD).



Figures II.4.1 and II.4.2: Proteins that interact with NFκB according to Ingenuity Pathway Analysis, a signalling factor that makes a strong contribution to the pro-inflammatory phenotype in CF.

3.5 Biochemical validation of differential expression by western blot analysis and nephelometry/immunoturbidimetry

Biochemical validation of NDKB (Mr 17kDa) abundance profile was performed by western blot analysis. Equal amounts of pooled samples were loaded in three independent gels to avoid inter-gel bias. The abundance levels were quantified in three independent experiments using densitometric analysis of immunoblots (Figure II.5A). These results unequivocally confirmed the increased abundance of NDKB in CF group, as identified by the spectral count analysis of the LC-MS/MS data.

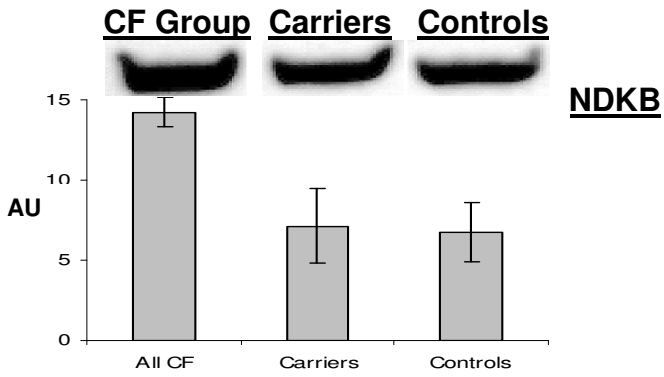


Figure II.5A: Biochemical validation of NDKB protein by western blot in pooled serum's depleted fraction from CF patients (n=35), carrier (n=21) and non-carrier controls (n=31), respectively (3 independent replicates, Progenesis PG200v2006 (NonLinear Dynamics was used for densitometry analysis).

Alpha-1-antitrypsin, complement C4 and ceruloplasmin concentrations were measured in individual serum samples by a nephelometric assay (Beckman Coulter Array® System). Significant differences for α-1-antitrypsin were observed when comparing group control with CF carriers and CF patients. No significant differences were observed when respiratory function was considered (Figure II.5B). As for complement C4, this protein was only identified by 2-DE and showed crescent expression when controls, CF carriers and CF patients were compared, pattern also observed in nephelometric determinations (Figure II.5C). Ceruloplasmin was identified in both 2-DE and LC-MS/MS and differences were only statistically significant when carriers and CF patients were compared although an increase tendency was observed when healthy individuals are compared with both mild and severe lung disease patients (Figure II.5D) in accordance with our findings.

Levels of apolipoprotein A-I (Apo A-I) and B (Apo B) were determined in individual serum samples using an immunoturbidimetric assay (Cobas®, Roche). Significant differences for Apo A-I were observed when comparing controls against all CF patients, patients with mild or severe lung disease and controls with severe lung disease patients (Figure II.5E). Recommend value (V_{rec}) for this protein in serum is >160mg/dL. Apo A-I levels in controls and carriers groups showed values above this threshold while all CF's group and severe CF's group revealed concentrations below this value. Apo B was determined as significantly different only when carriers and all CF patients were compared although a tendency of abundance's decrease is observed when controls are compared with CF patients (Figure II.5F). All the measurements for this protein were below the V_{rec} for Apo B in serum (<150mg/dL) probably due to long-term storage.

These results corroborate our observations from 2-DE and/or LC-MS/MS.

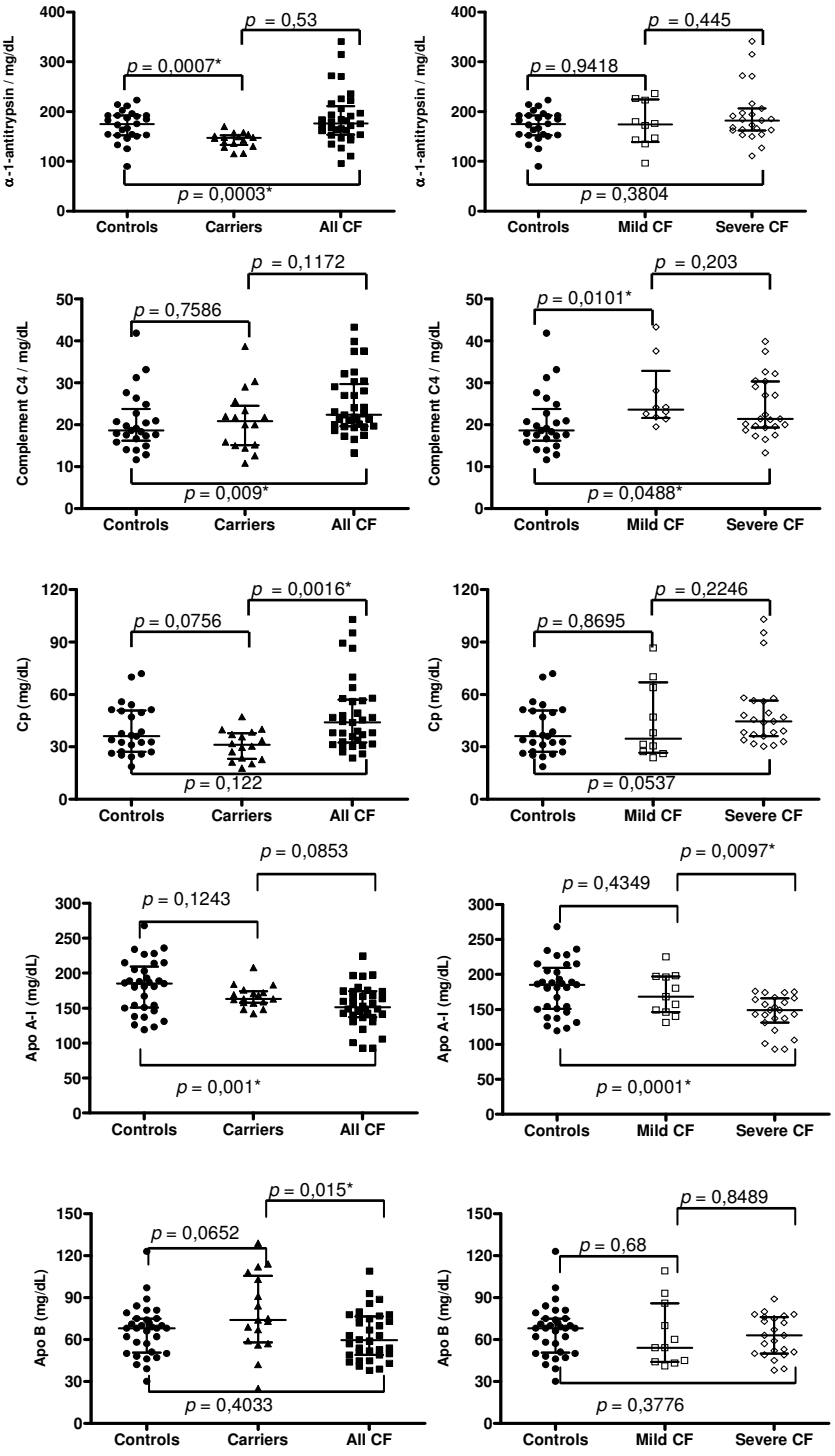


Figure II.5B-F: Biochemical validation of alpha-1-antitrypsin (**B**), complement C4 (**C**) and ceruloplasmin (**D**) by nephelometric assay using Beckman Coulter Array® Systems kits and apolipoprotein A-I (**E**) and apolipoprotein B (**F**) by immunoturbidimetric assay (Cobas®, Roche), respectively, in serum of the 87 individuals enrolled. Significantly different expression values between groups marked with *.

4. DISCUSSION AND CONCLUSIONS

Cystic Fibrosis is a lethal monogenic disease with no curative therapy and most patients die from progressive chronic lung dysfunction. The present study was conducted to compare serum proteome profiles of healthy non-CF individuals with carriers and CF patients and, among these, patients with mild or severe respiratory disease categorized by both clinical and spirometry-based criteria. Identification of new and more specific biomarkers for this pulmonary disease is of great importance to gain a better understanding of this multifactorial pathology as genetic tests alone cannot predict the heterology of phenotypes.

In this work, we used complementary 2-DE and LC-MS/MS-based approaches for global detection and relative quantification of medium and low abundant proteins with possible biological significance in CF. Biological processes such as tissue destruction and abnormal airway remodeling, protease/antiprotease imbalance or innate immune dysfunction were highlighted and are in accordance with recent findings [13]. Since serum was used as biological sample, the main pathways retrieved by Reactome were associated with hemostasis, lipids metabolism and platelet and complement cascade activation, as expected.

4.1 NFkB signaling pathway

The NFκB signaling pathway regulates many aspects of cellular activity, in stress-induced and inflammatory response but mainly in pathways of immune response [38]. Incorrect regulation of this transcription factor/pathway may cause inflammatory and autoimmune diseases, such as arthritis, and makes an important contribution to the pro-inflammatory phenotype of CF already described in lung epithelial cells of these patients [23,39]. Activation of NFκB is dependent of various stimuli including cytokines, stress inducers and/or viral/bacterial proteins. Using pathway analysis techniques, some NFκB-interacting proteins were identified including several members of the apolipoproteins family, complement system, CD180, BAG4, TNF ligand superfamily member 9, α-1-antitrypsin and regulator of G-protein signaling 3. Several other proteins such as hemopexin, vinculin, HSPA8, HSP70 and periaxin were also identified in this network, indicating possible deregulation in the defense mechanisms regulated by this protein complex. Regulation of this pathway is an important

phenomenon for preventing and/or attenuating the pro-inflammatory signaling properties caused by NFκB and improving clinical manifestations in patients.

4.2 Tissue remodeling and protease/antiprotease imbalance

It is well documented that pulmonary parenchyma destruction in CF is caused by persistent exposure to the action of proteolytic enzymes, such as neutrophilic elastase, cathepsin G or proteinase 3 [13,40], resulting in tissue destruction. Our results suggest a compensation of this destruction by increasing the levels of cytoskeleton proteins. Vinculin, for example, shows an up regulation by 20- and 10-fold in CF patients when compared with controls and carriers, respectively, suggesting a state of active tissue remodeling. No differences were found in CF patients with mild or severe lung disease, but its levels were increased in both conditions when compared with controls. Although not lung-specific, this protein is involved in cell adhesion and in the attachment of the actin-based microfilaments to the plasma membrane playing important roles in cell morphology and locomotion. Other examples of cytoskeleton proteins that are increased in abundance in CF patients are radixin (levels not detected in controls but 3 and 2 peptides identified in carriers and CF patients, respectively), tropomyosin-4 (3-fold), paladin (2-fold), myosin VIIa (2-fold) and dedicator to cytokinesis protein 1 (1.6-fold). Their involvement in cytoskeletal rearrangements, regulation of cell migration and adhesion and localization where actin remodeling takes place are strongly suggestive of processes at work to overcome the tissue destruction observed in CF patients [41].

Corroborating this tissue remodeling hypothesis is the observed decreased abundance level of α-1-antitrypsin. This protein was identified in both 2-DE and LC-MS/MS and found to be consistently decreased in abundance in carriers and CF patients but with no significant changes when the respiratory function was concerned. Only significant changes were observed in 2-DE spots when carriers were compared either with controls and CF patients thus confirming the previous result. α-1-antitrypsin is an antiprotease which main function is to inactivate neutrophils and inhibit neutrophil elastase, cathepsin G and proteinase 3, proteins already described as implicated in the tissue injury characteristic of CF lung disease [13,40]. Lungs of CF patients are constantly under an inflammation condition with recruitment and activation of neutrophils, promoting elastase's release and consequent destruction of pulmonary parenchyma [42]. Although considered an acute phase protein, α-1-antitrypsin has an

important function preventing this destruction acting as an antiprotease and anti-inflammatory agent in the lungs [43,44]. Several spots were identified as α-1-antitrypsin in 2-DE but no PTM was undoubtedly assigned by FindMod *in-silico* prediction probably due to other intracellular processing events and/or multiple PTMs *per* peptide that were not considered by the software. In CF, aerosol treatment with α-1 protease inhibitor has proven to restore the protease-antiprotease imbalance and suppress bacterial proliferation, reducing inflammation and tissue damage arising from chronic neutrophil activation. α-1-antitrypsin, along with 3 other serum proteins, was collectively evaluated and correctly discriminated lung cancer patients from healthy individuals [45-47].

Another group of differentially expressed proteins identified here was inter-α-trypsin inhibitors (ITI), namely heavy chains H1, H3 and H4, the latter identified by 2-DE and LC-MS/MS. Proteins of this family are believed to be present at medium to low abundance in serum and little is known about their function except the ability to bind hyaluronic acid in the extracellular matrix. Studies revealed that ITI levels in serum of septic patients were significantly lower than those found in healthy individuals [48] and that these proteins reduce the complement-dependent lung injury *in vivo* [49], promoting bronchial epithelial repair [50]. ITIH4 in particular was already identified as less abundant in COPD patients [51], although authors could not establish a linear correlation between levels of this protein and the pathology. In our study, ITI proteins were found to have decreased abundance in carriers and CF patients with similar expression level when respiratory function was considered. It is known that CF patients have recurrent lung infections and that several organs are affected and ITI levels might reflect this generalized state of inflammation as similarly to the septic patients analyzed. Also, and in accordance to Garantziotis *et al.* and Adair *et al.*, reduced levels of ITI might not inhibit complement activation so efficiently, perpetuating complement-induced lung injury, nor be able to restore the damaged lung epithelia.

4.3 Innate immune dysfunction

Differential abundance levels of the various components of the complement system were observed in serum of CF patients. Several components of this system, which acts in defense against infection and as anti-inflammatory agents, were identified:

increases in complement C5, C9 and C1s subcomponent, and decreases in C3, C3/C5 convertase and C6, suggesting that the innate immune system may be important in the development of the disease. Complement C3 was identified as differentially expressed in several gel spots (11 in total), 9 of them decreased in abundance. In two spots, PTMs were predicted *in-silico* as acetylation (spot 196) and lysine modification (spot 96). Decreased levels were also observed by LC-MS/MS. Complement C3 has been described as protective against the *Pseudomonas aeruginosa* in a murine model of pneumonia due to C3-mediated opsonization and phagocytosis of bacteria [52]. Also, complement C3 has been implicated in diseases that share features with CF, such as asthma [53], corroborating the obtained results. On the other hand, complement C5 was found to be increased in abundance in CF patients and slightly decreased when respiratory function was considered. Complement C5 is a central mediator in the complement activation cascade and mediator of local inflammatory processes, inducing vascular permeability and acting as chemoattractants to recruit immune cells toward sites of infection. It is known to be associated with direct tissue injury and immune deregulation [54,55] and patients with dysfunction of C5 display an increased propensity for severe recurrent infections [56,57]. Also, C5-deficient mice revealed higher mortality rates when infected with *P. aeruginosa* [58,59]. A relevant number of the CF patients analyzed here (60%) were infected with *P. aeruginosa* showing the relevance of the complement system in defense against pathogenic organisms in the lungs. Other studies however showed that C5 was significantly decreased in the BALF of CF individuals [13]. In the same study, complement C2 was undetectable in BALF by both MS and immunoblotting and its deficiency was associated with recurrent respiratory infections during childhood and adolescence. The present results corroborate this earlier observation as C2 was not identified here. However, complement C6 was identified as decreased in carriers and CF patients as well as in mild and severe lung disease groups. Association between C6 and lung function has already been described [60] where it has been shown that elevated levels of this component are present in human BAL after LPS exposure, although levels might differ from those in serum. Overall, the present results suggest that deregulation of the complement system may have a role in the pathophysiology of CF with concomitant inability to respond to chronic infections typical of CF lung disease or perpetuating tissue damage.

4.4 Oxidative stress and nutritional status

Several studies suggest that ceruloplasmin might have a protective function against free radicals-mediated damage in plasma [61] and lung [62] and that levels of this protein increase during inflammation, trauma and some malignancies. However, recent studies from our group [63] and others [64] showed no statistically significant differences in seric ceruloplasmin concentrations between CF patients and healthy individuals. Using a proteomics approach, we also did not observe significant differences in ceruloplasmin levels. A nephelometry-based assay showed differences only when carriers were compared with CF patients and an increase tendency when respiratory function of the enrolled patients is concerned, corroborating our findings. In spite of these results, antioxidant functions are undoubtedly attributed to ceruloplasmin. Although levels of this protein are not altered in CF patients, its activity and/or cellular location [65] might be relevant for the oxidative stress imbalance evidenced in these patients [66].

Another protein identified by both 2-DE and LC-MS/MS with established correlation with CF is the nutritional maker vitamin D-binding protein. This protein has shown to be decreased in carriers and CF patients, indicating inadequate and imbalanced nutritional status of these patients [67]. Unlike transferrin, transthyretin (TTR) and plasma retinol binding protein (PRBP), serum vitamin D-binding protein concentration seems to be independent of infection or inflammation, making it a potential indicator of the nutritional status in CF. These other three proteins were also found to be decreased in carriers and CF patients and, except for TTR, were only identified by 2-DE. Recent work from Chatterji *et al.* [68] showed inversed results with the levels of these proteins being up regulated in serum of a *c-myc* transgenic mice model. However, Roberts *et al.* [69] demonstrated that lower levels of TTR coincide with lower levels of retinol and PRBP and have been reported in ovarian cancer. Previous work from our group [70] has shown that retinoic acid metabolism, an important pathway in organ development and homeostasis including the lung, is altered in CF mutant mice lungs' recovered after naphthalene-induced injury, providing evidence that the decreased abundances of proteins identified here might be involved in tissue remodeling/repair. Lower levels of PRBP were also observed in COPD patients [51] and are in accordance with our own findings.

Apolipoproteins are another major proteins family that was identified to be deregulated in CF patients. These proteins function primarily as lipid-binding proteins involved in the regulation of the intravascular metabolism prior to their tissue uptake [68]. ApoA-I, ApoB and ApoE were identified as gradually decreased from non-CF carriers to healthy CF carrier individuals and from those to CF patients while reduced levels of ApoC2 and ApoF were observed in CF patients when compared with both healthy controls and carriers. Unaltered levels of ApoA4, ApoC3 and ApoL1 were also observed. Regulation of subclasses of apolipoproteins has been reported in some malignancies, namely ApoE which regulation was already described in colorectal [71], pancreatic [72] and lung cancer of c-myc transgenic mice [68], or ApoA-I which over expression has been implicated in breast and lung cancer [73,74]. Differential behavior is observed for this protein when its levels are compared in some respiratory diseases: higher levels of ApoA-I are found in asthma [75] and decreased expression has been reported in serum and BAL of adult smokers [76,77], with possible implications in COPD's development and progression. Moreover, ApoE was even recently considered a candidate biomarker for COPD [51], with increased levels in plasma of patients with this pathology due to its correlation with cardiovascular diseases that are high-prevalence comorbidities of COPD. ApoB is also associated with high risk for cardiovascular disease but its function as a barrier against infection was already described. Mice deficient in plasma apolipoprotein B are more susceptible to invasive *Staphylococcus aureus* infection [78], suggesting a role of essential innate defense barrier against this pathogen that also colonizes CF patients' lungs. Decreased levels of this protein in serum of CF patients as evidenced here might contribute to this susceptibility to aggressive *S. aureus* infection. Although specific role of the various Apo's remains to be elucidated in CF, altered levels of several apolipoproteins might be related with the oxidant/antioxidant imbalance verified in CF patients and poor nutrition status as well as new functions in infection by opportunistic pathogens.

4.5 Inflammation

Another lipoprotein identified as differentially abundant was serum amyloid A4 protein (SAA4). SAA4 was found to be down regulated in CF patients with no significant level changes in carriers, controls and when CF lung disease was accessed. Serum concentrations of this protein showed no correlation with other apolipoproteins,

such as ApoA-I, ApoB or ApoE, and that it is minimally induced by inflammation, when compared with SAA. Some authors have proposed SAA4 as a nutritional indicator with restrictions in a non-inflammatory state [79]. Although traditionally associated with Down's syndrome and Alzheimer's disease, correlation of SAA4 with CF has not been established. However, a member of its family, SAA, has already been associated with CF lung disease [80] and was also identified in this study only in the CF group, probably reflecting impairment of lung disease due to active *P. aeruginosa* infection.

Endothelin A receptor (ETA) was identified as lower abundant in CF patients in comparison with carrier and non-carrier controls, which have similar levels of abundance, although with low peptide counts in our samples (2 peptides in non-carrier and carrier controls and 1 in CF patients). Antagonists of this receptor have been implicated in attenuating pulmonary inflammation [81] and reducing smoke-related lung injury, caused by neutrophil recruitment and adhesion to pulmonary vascular endothelium [82] with putative interest in COPD. Involvement in CF is also expectable.

4.6 BiP and NDKB: Role in CF

All the proteins described so far have been demonstrated to be actively involved in several biological processes that are known to be altered in the pathogenesis of CF, namely defense response and immunity, complement activity, inflammation and wound repair. However, two of the proteins identified using the label free LC-MS/MS approach appear to have a central role in CF disease. Heat shock 70kDa protein 5 (HSPA5) and nucleoside diphosphate kinase B (NDKB) were both identified exclusively in the CF group with low counts revealing physiological low abundance of these proteins in serum. HSPA5, also known as glucose regulator protein 78kDa (GRP78) or BiP, is the main unfolded protein response (UPR) or stress sensor in cells [83]. BiP is also associated with trafficking, degradation and function of proteins including CFTR, playing an important role in protein processing [84]. UPR is induced in cell lines overexpressing the misfolded F508del-CFTR [85-87]. It is known that CF causes inflammation and cellular stress in several organs and therefore it is not surprising that BiP is elevated in serum of these patients. Its identification in CF circulation by this approach is an important finding of this study. BiP-associated protein was also identified only in the CF group by two peptides, emphasizing the role of the 'proteinaceous' environment in protein folding and maturation.

On the other hand, NDKB is a multifunctional enzyme which location and function are both complex and controversial and its identification only in serum of CF patients provides a strong hint of its importance in disease development, based on the functions described. In the cytosol, NDKB exists as a hexamer, whereas in the membrane monomers are described to bind G proteins and a variety of enzymes related to cellular energy production/utilization [88]. This protein was described as present in ovine [89], murine and human airway epithelium [90] and functions such as regulation of pancreatic secretion, neutrophil-mediated inflammation or energy generator have been described, suggesting that it might have an integrative role in the membrane channel function via ions as second messengers. In fact, this concept of NDKB as an ion sensor is of great relevance for the airway epithelium which undergoes rapid changes in tonicity (due to temperature or humidity) with potential to disturb the composition of the airway surface liquid bathing the cilia and consequently the control of mucus clearance in the lung. It has been shown that phosphorylation (and activity) of NDKB is differentially regulated by Cl^- and Na^+ [91], electrolytes which transport is altered in CF lung disease [92,93]. It was also proved that the activity of NDKB is abnormally low and dysfunctional in CF but restorable by retinoic acid [94,95]. Muimo *et al.* [89] showed that NDKB is mainly localized in an enriched apical membrane fraction from airway epithelium [96], localization that it shares with CFTR. This cellular locale is critical for ion transport, volume regulation and fluid secretion, processes all known to require protein kinases, Cl^- and GTP, all of which are imbalanced in CF. Other functions attributed to this protein account for cell growth, development and differentiation [89], tumor metastasis suppression [97] and promotion of epithelial cells proliferation [98]. In our study, NDKB was only identified in the CF group, probably indicating imbalanced electrolyte transport and active remodeling processes in affected epithelia. Its differential behavior and role in CF provide strong evidences that support the notion that NDKB might aid understanding of the pathogenesis of CF lung disease.

5. SUMMARY

In summary, the present investigation demonstrates that complementary gel-based and gel-free MS approaches are reliable tools to gain in-depth information on proteins actively involved in CF pathophysiology. The results reported here show that differences in the abundance level of several proteins are evident when comparing

serum of non-carrier controls and CF patients, revealing processes such as tissue remodeling, complement system dysfunction with consequent impairment on defense mechanisms and chronic inflammation, nutritional imbalance and *P. aeruginosa* colonization. Several studies confirmed these altered biological processes as implicated in the pathogenesis of CF, thereby corroborating the present findings. Future studies using a wider cohort of patients and orthogonal/complementary techniques will be beneficial to further verify and validate a serum proteomic signature to allow efficient discrimination of healthy individuals from CF patients and, ideally, from other chronic lung diseases. The present investigation represents a crucial a step-forward in characterization the complex changes in the serum proteome of CF patients with potential clinical importance. Although specific markers for CF were not found, several inflammatory-related proteins were highlighted contributing to elucidate the higher susceptibility of CF patients to some diseases as they seem to have a permanent activated inflammatory response.

REFERENCES

[1] Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R*, et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989; 245:1066-73.

[2] Ollero M, Brouillard F, Edelman A. Cystic fibrosis enters the proteomics scene: new answers to old questions. Proteomics 2006; 6:4084-99.

[3] Accurso FJ. Update in Cystic Fibrosis 2006. Am J Respir Crit Care Med 2007; 175:754-757.

[4] Cystic Fibrosis Foundation. Cystic Fibrosis Foundation, Patient Registry 2005 Annual Report, Bethesda, MD. 2004.

[5] Roxo-Rosa M, da Costa G, Luider TM, Scholte BJ, Coelho AV*, et al.* Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006; 6:2314-25.

[6] Penque D. Proteomic biomarker discovery for the monogenic disease cystic fibrosis. Expert Rev Proteomics 2007; 4:199-209.

[7] Merlo CA, Boyle, M.P. Modifier genes in cystic fibrosis lung disease. Journal Lab Clin Med 2003; 141:237-241.

[8] Gray RD, Macgregor G, Noble D, Imrie M, Dewar M*, et al.* Sputum Proteomics in Inflammatory and Suppurative Respiratory Diseases. Am J Respir Crit Care Med 2008.

[9] Sloane AJ, Lindner RA, Prasad SS, Sebastian LT, Pedersen SK*, et al.* Proteomic analysis of sputum from adults and children with cystic fibrosis and from control subjects. Am J Respir Crit Care Med 2005; 172:1416-26.

[10] Macgregor G, Gray RD, Hilliard TN, Imrie M, Boyd AC*, et al.* Biomarkers for cystic fibrosis lung disease: Application of SELDI-TOF mass spectrometry to BAL fluid. J Cyst Fibros 2008.

[11] Bai Y, Galetskiy D, Damoc E, Paschen C, Liu Z, *et al.* High resolution mass spectrometric alveolar proteomics: identification of surfactant protein SP-A and SP-D modifications in proteinosis and cystic fibrosis patients. Proteomics 2004; 4:2300-9.
[12] Sagel SD, Sontag, M.K, Accurso, F.J. Relationship between antimicrobial proteins and airway inflammation and infection in cystic fibrosis. Pediatr Pulmonol 2009; 44:402-409.
[13] Gharib S, Vaisar T, Aiken M, Park DR, Heinecke JW, *et al.* Mapping the lung proteome in cystic fibrosis. J Proteome Res 2009.
[14] Yocum A, Yu K, Oe T, Blair I. Effect of immunoaffinity depletion of human serum during proteomics investigations. J Proteome Res 2005; 4:1722-1731.
[15] Jacobs JM, Adkins JN, Qian WJ, Liu T, Shen Y, *et al.* Utilizing human blood plasma for proteomic biomarker discovery. J Proteome Res 2005; 4:1073-1085.
[16] Cho S, Lee E, Lee JC, Kim H, Park J, *et al.* Efficient prefractionation of low-abundance proteins in human plasma and construction of a two-dimensional map. Proteomics 2005; 5:3386-3396.
[17] Stalder D, Haeberli A, Heller M. Evaluation of reproducibility of protein identification results after multidimensional human serum protein separation. Proteomics 2008; 8:414-424.
[18] Hood BL, Lucas DA, Kim G, Chan K, Blonder J, *et al.* Quantitative analysis of low molecular weight proteome using ^{18}O stable isotope labeling in a lung tumor xenograft mouse model JASMS 2005; 16:1221-1230.
[19] Anderson NL, Anderson NG. History, character and diagnostic prospects. Mol Cell Proteomics 2002; 1:845-867.
[20] Zolotarjova N, Mrozinski P, Chen H, Martosella J. Combination of affinity depletion of abundant proteins and reversed-phase fractionation in proteomic analysis of human serum/plasma. J Chromatogr A 2008; 1198:332-338.
[21] Gong Y, Li X, Yang B, Ying W, Li D, *et al.* Different immunoaffinity fractionation strategies to characterize the human plasma proteome. J Proteome Res 2006; 5:1379-1387.
[22] Jmeian Y, Rassi Z. Micro-high-performance liquid chromatography pllatform for the depletion of high-abundance proteins and subsequent on-line concentration/capturing of medium and low-abundance proteins from serum. Application to profiling of protein exression in healthy and osteoarthritis sera by 2-D gel electrophoresis. Electrophoresis 2008; 29.
[23] Srivastava M, Eidelman O, Jozwik C, Paweletz C, Huang W, *et al.* Serum proteomic signature for cystic fibrosis using an antibody microarray platform. Molecular Genetics and Metabolism 2006; 87:303-310.
[24] Pedersen SK, Sloane AJ, Prasad SS, Sebastian LT, Lindner RA, *et al.* An immunoproteomic approach for identification of clinical biomarkers for monitoring disease: application to cystic fibrosis. Mol Cell Proteomics 2005; 4:1052-60.
[25] Wang G, Wu W, Zeng W, Chou C, Shen R. Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: Reproducibility, linearity and application with complex proteomes. J Proteome Res 2005; 5:1214-1223.
[26] Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendonza A, *et al.* Comparison of label-free methods for quantifying human proteins in shotgun proteomics. Mol Cell Proteomics 2005; 4:1487-1502.
[27] Johnson C, Buttler S, Konstan W, Morgan W, Wohl M. Factors influencing outcomes in cystic fibrosis: a center-based analysis. Chest 2003; 123:20-27.

[28] Pacheco P, Detection rate of CF causing CFTR mutations in Portugal. in: WHO, (Ed.), The Molecular Genetic Epidemiology of Cystic Fibrosis, Human Genetics Programme, Chronic Diseases and Health Promotion, 2004.
[29] Johnson C, Butler SM, Konstan MW, Morgan W, Wohl ME. Factors influencing outcomes in cystic fibrosis: a center-based analysis. Chest 2003; 123:20-7.
[30] Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM*, et al.* Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. Electrophoresis 2004; 25:1327-33.
[31] Shevchenko A, Wilm M, Vorm O, Mann M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. Anal Chem 1996; 68:850-858.
[32] Bensalem N, Masscheleyn S, Mozo J, Vallee B, Brouillard F*, et al.* High sensitivity identification of membrane proteins by MALDI TOF-MASS spectrometry using polystyrene beads. J Proteome Res 2007; 6:1595-602.
[33] Gomes-Alves P, Pesquita C, Couto F, Coelho AV, Penque D. Rescue F508del-CFTR by RXR-motif inactivation triggers proteome modulation associated with unfolded protein response. Biochim Biophys Acta - Proteins and Proteomics 2009.
[34] Matthews L, Gopinath G, Gillespie M, Caudy M, Croft D*, et al.* Reactome knowledgebase of biological pathways and processes. Nucleic Acids Res 2009.
[35] Liotta LA, Ferrari M, Petricoin EF. Written in blood. Nature 2003; 425:905.
[36] Jmeian Y, Rassi Z. Micro-high-performance liquid chromatography pllatform for the depletion of high-abundance proteins and subsequent on-line concentration/capturing of medium and low-abundance proteins from serum. Application to profiling of protein expression in healthy and osteoarthritis sera by 2-D gel electrophoresis. Electrophoresis 2008; 29:2801-11.
[37] Pesquita C, Faria D, Falcão AO, Lord P, Couto F. Semantic Similarity in Biomedical Ontologies. PloS Computational Biology 2009; 5.
[38] May B, Gosh. Rel/NF-kB and IKB proteins: an overview. Seminars in Cancer Biology 1997; 8:63-73.
[39] Tabary O, Boncoeur E, de Martin R, Pepperkok R, Clément A*, et al.* Calcium-dependent regulation of NFkB activation in cystic fibrosis airway epithelial cells. Cellular Signaling 2006; 18:652-660.
[40] Stoller JK, Aboussouaan LS. Alpha-1-trypsin deficiency. Lancet 2005; 365:2225-2236.
[41] Miller I, Eberrinni I, Gianazza E. Proteomics of lung physiopathology. Proteomics 2008; 8:5053-5073.
[42] Society AT, Society ER. Standards for the diagnosis and management of individuals with alpha-1-trypsin deficiency. Am J Resp Crit Care Med 2003; 168:818-900.
[43] Stockley RA, Bayley DL, Unsal I, Dowson LJ. The effect of augmentation therapy on bronchial inflammation in alpha-1-antitrypsin deficiency. Am J Resp Crit Care Med 2002; 165:1494-1498.
[44] Camelier AA, Winter DH, Jardim JR, Barboza CE, Cukier A*, et al.* Alpha-1-antitrypsin deficiency: diagnosis and treatment. J Bras Pneumo 2008; 34:514-527.
[45] Patz EFJ, Campa MJ, Gottlin EB, Kusmartseva I. Panel of serum biomarkers for the diagnosis of lung cancer. J Clinical Oncology 2007; 25:5578-7783.
[46] Taguchi F, Solomon B, Gregorc V, Roder H, Gray R*, et al.* Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment

with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. J Natl Cancer Inst 2007; 99:838-846.
[47] O'Byrne KJ, Danson S, Dunlop D, Botwood N, Taguchi F*, et al.* Combination therapy with gefitinib and rofecoxib in patients with platinum-pretreated relapsed non-small-cell lung cancer. J Clin Oncology 2007; 25:3266-3273.
[48] Lim YP, Bendelja K, Opal SM, Siryaporn E, Hixson DC*, et al.* Correlation between mortality and the levels of inter-alpha-trypsin inhibitors in plasma of proteins with severe sepsis. J Infectious Disease 2003; 188:919-926.
[49] Garantziotis S, Holligsworth JW, Ghanayem RB, Timberlake S, Zhou L*, et al.* Inter-alpha-trypsin inhibitor attenuates complement anctivation and complement-induced lung injury. J Immunol 2007; 179:4187-4192.
[50] Adair JE, Stober VP, Sobhany M, Zhou L, Roberts JD*, et al.* Inter-alpha-trypsin inhibitor promotes bronchial epithelial repair after injury through vitronectin binding. J Biol Chem 2009.
[51] Bandow JE, Baker JD, Berth M, Painter C, Sepulveda OJ*, et al.* Improved image analysis workflow for 2-D gels enables large-scale 2-D-based proteomics studies - COPD biomarker discovery study. Proteomics 2008; 8:3030-3041.
[52] Mueller-Ortiz SL, Drouin SM, Wetsel RA. The alternative activation pathway and complement component C3 are critical for a protective immune response against Pseudomonas aeruginosa in a murine model of pneumonia. Infection and Immunity 2004; 72:2899-2906.
[53] Rhim T, Choi YS, Nam BY, Uh ST, Park JS*, et al.* Plasma protein profiles in early asthmatic responses to inhalation allergen challenge. Allergy 2009; 64:47-54.
[54] Ward PA. The dark side of C5a in sepsis. Nat Rev Immunol 2004; 4:133-142.
[55] Huber-Lang MS, Younkin EM, Sarma JV, McGuire SR, Lu TK*, et al.* Complement-induced impairment of innate immunity during sepsis. J Immunol 2002; 169:3223-3231.
[56] Pfarr N, Prawitt D, Kirschfink M, Schroff C, Knuf M*, et al.* Linking C5 deficiency to an exon splicing enhancer mutation. J Immunol 2005; 174.
[57] Guo RF, Ward PA. Role of C5a in inflammatory processes. Annual Reviews in Immunology 2005; 23:821-852.
[58] Cerquetti MC, Sordelli DO, Bellanti JA, Hooke AM. Lung defenses against Pseudomonas aeruginosa in C5-deficient mice with different genetic backgrounds. Infect Immun 1986; 52:853-857.
[59] Larsen GL, Mitchel BC, Harper TB, Henson PM. The pulmonary response of C5 sufficient and deficient mice to Pseudomonas aeruginosa. Am Rev Respiratory Disease 1982; 126:306-311.
[60] Bolger MS, Ross DS, Jiang H, Frank MM, Ghio AJ*, et al.* Complement levels and activity in the normal and LPS-induced injury. Am J Physiol Lung Cell Mol Physiol 2007; 292:748-759.
[61] Halliwell B, Gutteridge J. Oxygen free radicals and iron in relation to Biology and Medicine: some problems and concepts. Arch Biochem Biophys 1986; 246:501-514.
[62] Fleming R, Whitman I, Gitlin J. Induction of ceruloplasmin gene expression in rat lung during inflammation and hypoxia. American Journal of Physiology 1991; 260:L60-L74.
[63] Guardiano M, Cystic Fibrosis: Genotype/Phenotype correlation and oxidative stress parameters in a Portuguese subpopulation, Medical School, University of Porto, Porto, 2004, pp. 109.

[64] Percival SS, Kauwell GPA, Bowser E, Wagner M. Altered copper status in adult men with cystic fibrosis. Journal of American Colleage of Nutrition 1999; 18:614-619.
[65] Banha J, Marques L, Oliveira R, Martins Mde F, Paixao E*, et al.* Ceruloplasmin expression by human peripheral blood lymphocytes: a new link between immunity and iron metabolism. Free Radic Biol Med 2008; 44:483-92.
[66] Bowler RP, Crapo JD. Oxidative stress in airways: is there a role for extracellular superoxide dismutase. Am J Respir Crit Care Med 2002; 166:S38-S43.
[67] Speeckaert MM, Wehlou C, Vadewalle S, Taes YE, Robberecht E*, et al.* Vitamin D binding protein, a new nutritional marker in cystic fibrosis patients. Clin Chem Lab Med 2008; 46:365-370.
[68] Chatterji B, Borlak J. A 2-DE MALDI-TOF study to identify disease regulated serum proteins in lung cancer of c-myc transgenic mice. Proteomics 2009; 9:1044-1056.
[69] Roberts D, Williams SJ, Cvetkovic D, Weinsteins JK, Godwin AK*, et al.* Decreased expression of retinol binding protein is associated with malignant transformtaion of the ovarian surface epithelium. DNA Cell Biology 2002; 21:11-19.
[70] Carvalho-Oliveira I, Charro NA, Aarbiou J, Buijs-Offerman R, Wilke M*, et al.* Proteomics analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009.
[71] Li ZG, Zhao L, Liu L, Ding YQ. Monitoring changes of serum markers in metastatic colorectal carcinoma model. Zhonghua Ming Li Xue Za Zhi 2007; 36:48-52.
[72] Gronborg M, Kristiansen TZ, Iwahori A, Chang R, Reddy R*, et al.* Biomarker discovery from pancreatic cancer secretome using a differencial proteomic approach. Mol Cell Proteomics 2006; 5:157-171.
[73] Chen JH, Anderson MP, Misek DE, Simeone DM, Lubman, D.M. Characterization of apolipoprotein and apolipoprotein precursors in pancreatic cancer serum samples via two-dimensional liquid chromatoghraphy and mass spectrometry. J Chromatogr A 2007; 1162:117-125.
[74] Fuji K, Kondo T, Yokoo H, Yamada T, Iwatsuki K*, et al.* Proteomic study of human hepatocellular carcinoma using two-dimensional difference gel electrophoresis with saturation cysteine dye. Proteomics 2005; 5:1411-1422.
[75] Nagel G, Weiland SK, Rapp K, Link B, Zoellner I*, et al.* Association of apolipoproteins with symptoms of asthma and atopy among schoolchildren. Int Arch Allergy Immunol 2009; 49:259-266.
[76] Craig WY, Palomaki GE, Haddow JE. Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data. BMJ 1989; 298:784-788.
[77] Lindahl M, Stahlbom B, Tagesson C. Two-dimensional gel electrophoresis of nasal and bronchoalveolar lavage fluids after occupational exposure. Electrophoresis 1995; 16:1199-1204.
[78] Peterson MM, Mack JL, Hall PR, Alsup AA, Alexander SM*, et al.* Apolipoprotein B is an innate barrier against invasive Staphylococcus aureus infection. Cell Host Microbe 2008; 4:555-66.
[79] Yamada T, Miyake N, Itoh K, Igari J. Further characterization of serum amyloid A4 as a minor acute phase reactant and a possible nutritional marker. Clin Chem Lab Med 2001; 39:7-10.

[80] Smith JW, Colombo JL, McDonald TL. Comparison of serum amyloid A and C-reactive protein as indicators of lung inflammation in corticosteroid treated and non-corticosteroid treated cystic fibrosis patients. J Clin Lab Anal 1992; 6:219-224.
[81] Kalk P, Senf P, Deja M, Petersen B, Busch T, *et al.* Inhalation of an endothelin receptor A antagonist attenuates pulmonary inflammation in experimental acute lung injury. Can J Physiol Pharmacol 2008; 86:511-515.
[82] Bahvsar T, Liu XJ, Patel H, Stephani R, Cantor JO. Preferential recruitment of neutrophils by endothelin-1 in acute lung inflammation induced by lipopolysaccharide or cigarette smoking. International Journal of COPD 2008; 3:477-481.
[83] Zhang K, Kaufman RJ. Signaling the unfolded protein response from the endoplasmic reticulum. J Biol Chem 2004; 279:25935-8.
[84] Wang X, Venable J, LaPointe P, Hutt DM, Koulov AV, *et al.* Hsp90 cochaperone Aha1 downregulation rescues misfolding of CFTR in cystic fibrosis. Cell 2006; 127:803-15.
[85] Gomes-Alves P, Couto F, Pesquita C, Coelho AV, Penque D. Rescue of F508del-CFTR by RXR motif inactivation triggers proteome modulation associated with the unfolded protein response. Biochim Biophys Acta 2010; 1804:856-865.
[86] Gomes-Alves P, Penque D. Highthroughput proteomics data uncover possible key players in F508del-CFTR processing and trafficking. Expert Rev Proteomics 2010.
[87] Jurkuvenaite A, Chen L, Bartoszewski R, Goldstein R, Bebok Z, *et al.* Functional stability of rescued delta F508 cystic fibrosis transmembrane conductance regulator in airway epithelial cells. Am J Respir Cell Mol Biol 2010; 42:363-72.
[88] Engel M, Seifer M, Theisinger B, Seyfert U, Welter C. Glyceraldehyde-3-phosphate dehydrogenase and nm23-H1/nucleoside diphosphate kinase A. Two old enzymes combine for the novel nm23 protein phosphotransferase function. J Biol Chem 1998; 273:20058-20065.
[89] Muimo R, Banner SJ, Marshall LJ, Mehta A. Nucleoside diphosphate kinase and chloride-sensitive protein phosphorylation in apival membranes from ovine airway epithelium. Am J Respir Cell Mol Biol 1998; 18:270-278.
[90] Treharne KJ, Riemen CE, Marshall LJ, Muimo R, Mehta A. Nucleoside diphosphate kinase - a component of the sodium and chloride-sensitive phosphorylation cascade in human and murine airway epithelium. Eur J Physiol 2001; 443:S-97-S102.
[91] Marshall LJ, Muimo R, Riemen CE, Mehta A. Na+ and K+ regulate the phosphorylation state of nucleoside diphosphate kinase in human airway epithelium. Am j Physiol 1999; 276:C109-C119.
[92] Mall M, Grubb BR, Harkema JR, O'Neal WK, Boucher RC. Increased airway epithelial Na+ absorption produces cystic fibrosis-like lung disease in mice. Nat Med 2004; 10:487-93.
[93] Riordan JR. Assembly of functional CFTR chloride channels. Annu Rev Physiol 2005; 67:701-18.
[94] Treharne KJ, Mehta A, Muimo R. NDKB, a protein kinase defective in CF, binds to and is regulated by AMPK, a CFTR-associated protein. Pediatr Pulmonol 2001; S22:240.
[95] Treharne KJ, Best OG, Mehta A. Transglutaminase 2 and nucleoside diphosphate kinase activity are correlated in epithelial membranes and are abnormal in cystic fibrosis. FEBS Lett 2009; 583:2789-2792.

[96] Collins FS. Cystic fibrosis: molecular biology and therapeutic implications. Science 1992; 256:774-9.
[97] Schultz DM, Böllner C, Thomas G, Atkinson M, Esposito I*, et al.* Identification of differentially expressed proteins in triple-negative breast carcinomas using DIGE and mass spectrometry. J Proteome Res 2009; 8:3430-3438.
[98] Desvignes T, Pontarotti P, Fauvel C, Bobe J. Nme protein family evolutionary history, a vertebrate perspective. BMC Evol Biol 2009; 9.

Chapter III

Changes in the proteome of CF's RBC membranes reveal altered RBC's scaffold and shape

Charro N[1,2], Simões T[1,2], Faria D[5], Pacheco P[3], Lopes C[4], Azevedo P[4], Bugalho de Almeida A[4], Couto FM[5], Chan KC[2], Isaaq HJ[2], Waybright T[2], Veenstra TD[2], Blonder J[2], Penque D[1]

[1]*Laboratory of Proteomics, Dept. of Genetics, INSA Dr. Ricardo Jorge, Lisbon, Portugal*
[2]*Laboratory of Proteomics and Analytical Technologies, SAIC-Inc., NCI at Frederick, MD, USA.*
[3]*Laboratory of Molecular Biology, Dept. of Genetics, INSA Dr. Ricardo Jorge, Lisbon, Portugal.*
[4]*Clínica Universitária de Pneumologia, Hospital de Santa Maria, Faculdade de Medicina da Universidade de Lisboa.*
[5]*Dept. of Informatics, Faculdade de Ciências da Universidade de Lisboa, Portugal.*

ABSTRACT

The influence of CF on red blood cells (RBC) normal physiology is not yet fully understood. Wild-type CFTR facilitates a deformation-induced ATP release in RBC that activates NO synthesis causing vessel dilation, condition that is absent in CF patients' RBC and might cause pulmonary hypertension. Therefore, investigating changes in RBC proteome can be a reliable tool to better understand CF.

Methanol-facilitated proteins' extraction combined with label-free LC-MS/MS was used to identify differences in abundance of membrane proteins from RBC isolated from peripheral blood of CF patients with mild or severe pulmonary disease in comparison with healthy CF-carriers and non-CF individuals. From a list of 520 proteins, 178 of them were further investigated in terms of biological function and relevance in the disease. Proteins involved in cytoskeleton rearrangements (CTSG, FLOT1, FLOT2, STOM, spectrins, ANK1, ACTB, GYPA, GYPC, EPB41, EPB42), antioxidant (CAT, MPO, EPX, PRDX1, PRDX2) and transport activities (SLC4A1, SLC4A2, SLC2A1, SLC2A4, SLC16A1, SLC2A14, HBD, HBG1, HBA1, HBB) were retrieved as highly significant in processes taking place in CF, corroborating previous studies and emphasizing their relevance in the context of CF.

Proteomic investigation of mRBC in the context of CF provided extra understanding of the complex changes that characterize this monogenic disease.

Keywords: Cystic Fibrosis; Red Blood cells; Two-Dimensional Liquid Chromatography coupled to Mass Spectrometry; Functional Analysis

1. BACKGROUND

Mature erythrocytes, or red blood cells (RBC), are 8-μm biconcave disks bounded by a plasma membrane that constitute the most abundant solid component of blood. These enucleated cells are the end stage of a process of maturation that starts in the bone marrow till the mature forms that have a life span of around 4 months in circulation, delivering oxygen to tissues in exchange of toxic carbon dioxide. The RBC's membrane (mRBC) is formed by almost equally amounts of lipids and proteins, creating a network constituted mainly by glycophorin A, band 3 protein, spectrin, actin, ankyrin, among others, and phospholipids. This arrangement is responsible for cellular deformability and flexibility needed for performing their function [1].

The proteome of RBC is however much more complex than initially thought despite its lack of intracellular organelles. Several works have emerged in the past few years with more and more complete characterization of the mRBC proteome [2-8].Our own group has analysed the mRBC proteome in the context of Chronic Obstructive Pulmonary Disease (COPD) and has confidently identified over 1000 proteins *(data not published)* corroborating the proteome's complexity of an apparent simple cell.

In Cystic Fibrosis (CF), absence or defective CF transmembrane conductance regulator (CFTR) on the exocrine epithelial tissues results in the failure of ionic and water homeostasis leading to excessive mucous production that causes chronic lung disease, pancreatic insufficiency, elevated concentrations of Cl^- in sweat, neonatal *meconium ileus*, infertility, among other symptoms [9,10]. However, it is not evident whether disturbances at the systemic and, most specifically, at the RBC level can influence disease phenotype. Mild haemolytic anaemia has been reported in some CF patients though its causes are attributed to reduced antioxidant defense [11] with consequent increase in elevated mRBC lipid peroxidation [12].

The controversy on the presence of CFTR in the RBC extends back to the early 1990's decade being undoubtedly unravelled that CFTR is present not only in epithelial cells but also in RBC in 1998 [13]. Since then, several reports confirmed this finding [14-16]. The most solid study so far was performed by Lange *et al.* where it was possible to count the number of CFTR molecules in RBC and observed a reduction of CFTR by 4 fold in CF patients when compared with healthy individuals [17].

It has been suggested that CFTR facilitates a deformation-induced ATP release in RBC while sneaking out the capillaries [13]. Although the levels of ATP in plasma of CF patients appear to be significantly higher than in non-CF [18], the same is not

verified in RBC with equal concentrations in both groups [13]. Under native CFTR, RBC release ATP in response to mechanical deformation in a stimulus-dependent fashion [13] that should activate nitric oxide (NO) synthesis in neighbouring endothelial cells causing vessel dilation [19-21]. However, erythrocytes from CF patients failed to stimulate NO synthesis in isolated rabbit lungs [22] strengthening the view that reduced deformation-triggered ATP release leads to an increase in vascular smooth muscle tone potentially causing pulmonary hypertension observed in some of these patients.

The multifunctional RBC with their roles in oxygenation, oxidative stress and vessel dilation can be a reliable biological sample to investigate differentially expressed proteins with possible interest in CF.

In the present work, we aim to identify differentially expressed proteins in mRBC of CF patients with mild or severe pulmonary disease in comparison with healthy individuals by high-throughput label-free 2D-LC-MS/MS and explore their relevance in the context of the disease and, in particular, the CF's lung disease severity.

2. MATERIAL AND METHODS

2.1Patients and Sample Collection

The present study was approved by the Ethics Committee of INSA Dr. Ricardo Jorge and Hospital de Santa Maria (HSM) – Lisbon, Portugal.

Thirty five CF patients bearing heterogeneous mutations in the *CFTR* gene under clinical surveillance in Pulmonology's Division of HSM-Lisbon were enrolled for this study. All CF patients were free of acute phase of inflammation and under standard respiratory treatments, namely nebulized recombinant human dornase alpha (rhDNAse), tobramycin and azithromycin to reduce sputum viscoelasticity, control respiratory infection and exacerbation and inhibit bacterial proliferation, respectively. CF patients were categorized into mild or severe according to [23] based on their forced expiratory volume in 1 second (FEV_1) and clinical criteria (Table III.1). The genotype heterogeneity observed in our subpopulation praises to identify significant differences between CF and non-CF groups without being mutation-restricted. Pooled samples from age- and gender-matched healthy individuals (n=32) and healthy carriers of one mutated *CFTR* allele (n=20) were used for comparison purposes. All healthy individuals were previously genotyped for the most common mutations in the Portuguese population [24]

to confirm that they were non-CF carriers and clinically evaluated (questionnaires, chest auscultation and spirometry) to address respiratory function ensuring a $FEV_1 \geq 80\%$ of vital capacity.

Table III.1: Cystic Fibrosis (CF) patients and Carriers demographics and stratification into disease severity based on clinical criteria and pulmonary function accessed by spirometry [23]. Healthy individuals were chosen to match similar demographic conditions as the CF patients.

	Group				
	Controls	**CF Patients**			**Carriers**
		Total	**Mild**	**Severe**	
n	32	35	11	24	20
Age (mean ± SD)	(30 ± 8) y	(27 ± 8) y	(27 ± 9) y	(27 ± 8) y	(45 ± 11) y
Gender	20 F 12 M	20 F 15 M	7 F 4 M	13 F 11 M	12 F 8 M
Genotype ΔF/ΔF (%)	NA	7 (20%)	0 (0%)	7 (20%)	NA
Genotype ΔF (%)	NA	22 (63%)	8 (73%)	14 (58%)	10 (48%)
FVC % (mean ± SD)	93 ± 11	77 ± 24	89 ± 23	71 ± 23	---
FEV_1 % (mean ± SD)	91 ± 12	60 ± 27	80 ± 24	51 ± 24	---
Colonization P.aeruginosa (%)	---	60	45	67	---
BMI (mean ± SD)	---	21.1 ± 3.5	22.9 ± 3.6	20.3 ± 3.2	---

2.2 RBC Isolation from Peripheral Blood

All the procedures were performed on ice (except were stated) and in the same day of blood collection to avoid cell degradation.

Peripheral blood was collected by veinipuncture into a 4.9ml blood-collection tube containing 1.6mg EDTA/ml blood and immediately kept on ice for a ~2 hours period (never exceeding 4 hours). Whole blood was centrifuged at 3000 x *g* for 5 min at 4 ºC to separate the solid and the liquid fractions. Both plasma and the buffy coat (white layer on top of RBC containing mainly leukocytes and thrombocytes) were carefully discarded and an additional centrifugation for another 2 min to remove evident traces of plasma and remaining buffy coat was performed. Enriched fraction in RBC was washed 3 times with 3 volumes of cold isotonic buffer [0.9% NaCl (w/v) solution, pH 8.0] by careful inversion of the tube to homogenize the mixture and avoid cell disruption. After

each wash, samples were centrifuged at 3000 x *g* for 5 min at 4 ºC to discard any left plasma and buffy coat layer.

2.3 RBC Microsomal's Fraction Enrichment

Preparation and isolation of mRBC was achieved according to [13,25]. Briefly, RBC were incubated in 2 volumes of cold hypotonic phosphate buffer [5mM (w/v)], pH 7.4 supplemented with protease inhibitors (SIGMA, P8340-5mL) for 30 min on ice for cell lysis. Frequent homogenization was performed by inversion of the tube. Fifteen volumes of hypotonic phosphate buffer were added to the mixture and centrifuged at 4500 x *g* for 10 min at 4 ºC. An aliquot of the supernatant (intracellular content) was saved, the remaining discarded and pellet was washed three times with phosphate buffer by centrifuging at 27,300 x g for 10 min at 4 ºC. A final wash in phosphate buffer at 25,000 x *g* for 5 min at 4 ºC was performed to guarantee the isolation of hemoglobin-free RBC ghost that were further stored at -80 ºC.

Individual samples from the above mentioned groups were pooled together to emphasize proteomic differences among the groups studied while eliminating potential individual contributions. The pooling strategy has been widely used in profiling studies due to its power in evidencing the significant differences between groups/conditions and reducing the financial and temporal stress in analyzing a large number of samples. Pooled samples were transferred to 2mL siliconized tubes and incubated with 50mM (w/v) NH_4HCO_3 pH 7.4. Vigorous shacking and sonication (5 x 15'; 30' intervals) was performed to aid membrane disruption before centrifugation at 100,000 x *g* for 1 hour at 4 ºC. Supernatant was discarded, the pellet resuspended in 100mM (w/v) Na_2CO_3 and incubated for 2h at 4ºC on a rotisserie followed by ultracentrifugation at 100,000 x *g* for 90 min at 4 ºC. Resulting pellet was washed twice with ddH_2O to obtain the membrane-enriched fractions using previously mentioned centrifugal conditions. Protein content was established by BCA™ Protein Assay Kit (Pierce). A total of 100 µg of lyophilized proteins from membrane-enriched fraction was digested overnight with sequence-grade modified trypsin (Promega, Madison, WI, USA) at an enzyme:protein ratio of 1:20 in 50mM NH_4HCO_3/60% (v/v) methanol [26]. Tryptic peptides were desalted by SPE (3M™ Empore™ High Performance Extraction Disk Cartridges) according to manufacturer's description and lyophilized to dryness prior to 2D-LC-MS/MS analysis.

2.4 2D-LC-MS/MS, Data Processing and Bioinformatic Analysis

Desalted lyophilized peptides were solubilized in 45% (v/v) ACN/0.1% (v/v) FA to achieve a concentration of approximately 0.5-1 µg/µl before separation and analysis by label free 2D-LC-MS/MS under standard conditions [26,27].

The CID spectra were analyzed using SEQUEST^{TM} operating on a Beowulf 18-node parallel virtual machine cluster computer (ThermoElectron, Thermo Fisher Scientific, Waltham, MA, USA) using a UniProt non-redundant human proteome database (http://www.expasy.org, 03/2008 release). Only peptides with conventional tryptic termini (allowing for up to two internal missed cleavages) possessing delta-correlation scores (ΔCn) >0.1 and charge state-dependent cross correlation (Xcorr) criteria of >2.1 for $[M+H]^{1+}$, >2.3 for $[M+H]^{2+}$ and >3.5 for $[M+H]^{3+}$were considered as legitimate identifications. A final list of the identified proteins with the corresponding number of unique (non-redundant) peptides (UPs) and total counts for those peptides (TPs) was obtained. Protein abundances were calculated by comparing total spectral counting values across analysed conditions and proteins were considered differentially expressed when showing ± 1.5 fold difference in abundance.

Sets of differentially expressed proteins were further characterized in terms of chemical properties [28-30] [determination of putative transmembrane domains (TMD), http://www.cbs.dtu.dk/services/TMHMM, and grand average hydropath value index (GRAVY), http://www.geneinfinity.org/sms_proteingravy.html] and biological information [Protein Information Knowledge Extractor (PIKE) [31], http://proteo.cnb.uam.es:8080/pike/, Ingenuity Pathway Analysis (IPA), www.ingenuity.com and ProteinOn [32], retrieving important functions, pathways and significantly overrepresented GO terms in the datasets.

3. RESULTS AND DISCUSSION

Several studies have provided lists of proteins characterizing the proteome of RBC proving that it is much more complex than initially thought [2-8]. However, the comprehensive RBC proteome characterization in the CF context is, to our knowledge, not yet investigated.

By using high-throughput label-free 2D-LC-MS/MS-based proteomics approach, we aimed to identify differences in the mRBC proteome profile that could be associated with CF and/or with CF lung disease severity. mRBC isolated from CF patients with severe (n=22) and mild (n=11) lung disease and from controls, health CF-Carriers (n=20) and health non-CF Carriers (n=32) were investigated.

From an initial list of 520 proteins identified in all groups considered (Table III.1.1SD), only proteins identified by at least 2 peptides (TPs) in any independent group were considered to have extra confidence in the identification, shortening the list to 178 proteins (Table III.1.2SD). The false discovery rate (FDR) calculated by searching the data against a decoy database was 4.3%. The relative abundance of proteins was calculated by comparing total spectral counts from the same protein across conditions and a threshold of ± 1.5 fold difference in abundance was set to consider them differentially expressed.

3.1 General categorization of the proteins identified in mRBC

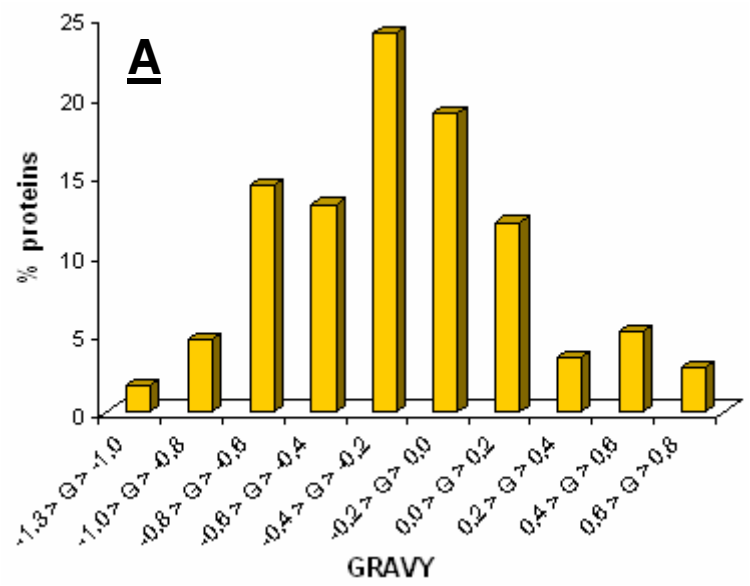
The 178 proteins considered in our study were further categorized in terms of hydrophobicity (GRAVY), number of predicted TMD and subcellular location to corroborate improved sample preparation methodology and effective analysis of the mRBC proteome.

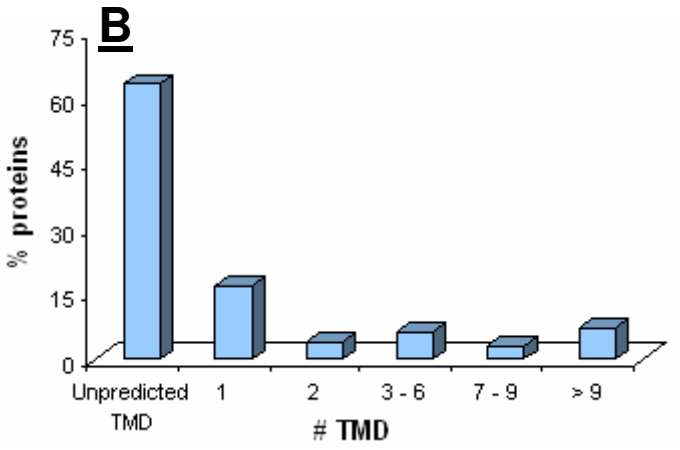
GRAVY is an overall index of protein solubility that takes in account the hydrophobicity values for each of the aminoacids in the protein sequence normalized to the protein's length. Positive GRAVY values are considered a reliable marker of the hydrophobic nature of a protein and a valid indicator of its membrane involvement. In our dataset, 23.4% of proteins present positive GRAVY values (GRAVY $\geq$ 0), proportion that rises to 42.3% when GRAVY $\geq$ -0.2 (Figure III.1A).

Mapping for putative TMD relied on the analysis of protein's sequences to obtain the correspondent hydropathy plots. Regarding this parameter, 36.9% of the identified specimens have at least one TMD (Figure III.1B). A significant percentage evidenced more than 9 TMD (7.1%) with protein FAM38A, a multipass endoplasmic reticulum membrane protein, predicted to have 23 TMD.

Subcellular location showed a successful enrichment in membrane proteins by retrieving 46% of proteins allocated to the membrane compartment (both organelle and

plasma membrane) (Figure III.1C). In addition, several other proteins were categorized as cytoskeletal and could be membrane-associated or membrane-anchored.





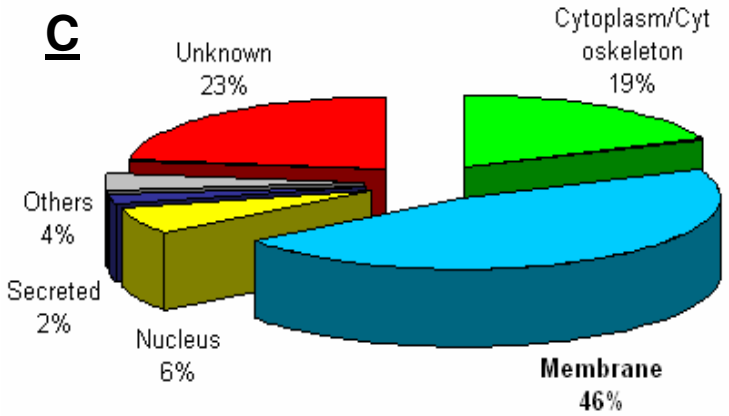


Figure III.1: Hydrophobicities (GRAVY values; Gene Infinity tools; http://www.geneinfinity.org/sms_proteingravy.html) (**III.A**), predicted transmembrane domains (TMDs; TMHMM software; http://www.cbs.dtu.dk/services/TMHMM/) (**III.B**) and primary location (based on information retrieved by PIKE from UniProt) (**III.C**) of proteins identified in mRBC.

Overall, our results indicate that the enrichment in membrane proteins was successfully achieved.

3.1.1 Identification of CFTR in mRBC

Sprague *et al.* in 1998 showed for the first time CFTR activity in RBC that contributes to deformation-induced ATP release is these cells [13]. Eight years after, Lange *et al.* (2006) by using atomic force microscopy combined with quantum-dot-labeled anti-CFTR antibody were able to count the number of CFTR molecules in mRBC that were reduced in CF ΔF homozygous patients in comparison with healthy subjects [17]. Despite of these studies, the unequivocally expression of CFTR in RBC remains to be demonstrated by MS-based approach. In our MS datasets, CFTR was identified by only one peptide in mRBC from ΔF/ΔF patient's group (Table III.1.1SD). CFTR's isolation and identification by MS is challenging due to its high hydrophobicity, glycosilation and relative low expression levels in mRBC [33]. The fact, however, of CFTR being identified exclusively in the ΔF/ΔF patients is nevertheless intriguing. One possible explanation for that is the intracellular accumulation of ΔF-CFTR due to folding defects and recognition by the cell's quality control to overcome such flaw or send the protein to degradation [34]. Such accumulation might result in increasing the abundance of CFTR is cells from this group overcoming the detection limit of the MS thus making possible its detection. In spite of all of this, it is however noteworthy that we were able to identify CFTR by MS in this biological sample by a 'shotgun' approach, confirming unequivocally the existence of this protein in mRBC.

3.2 Differentially expressed proteins as consequence of CF

3.2.1 Functional Analysis and Characterization

Although a broad characterization was achieved by calculating GRAVY, TMD and investigating the primary subcellular location of the identified proteins, the main focus of a proteomics study is to further explore the differences in expression between two (or more) conditions. Functional characterization of the differentially expressed proteins was performed by ProteinOn and Ingenuity Pathway Analysis (IPA).

ProteinOn is a web tool that calculates GO-based protein semantic similarity and implements a preliminary weighting factor (info content), which increases the specificity of existing semantic similarity measures, and a probabilistic score (*e*-value) for representativeness of a GO term within a set of proteins [32]. With this unified and structured vocabulary, integration of knowledge from the sequence level to the system level is facilitated. Table III.2 compiles some of the most significant molecular functions retrieved by the analysis of differentially expressed proteins' sets. The complete list of all molecular functions retrieved is provided in Table III.2SD.

From the analysis of these terms, it is evident that structural organization of the cytoskeleton, transport activity and oxidant/antioxidant status account for the most significant molecular functions in which the differentially expressed proteins participate in, confirmed by particularly low *e*-values. In fact, transport activity are spread throughout several categories and substrates that include oxygen (HBD, HBG1, HBA1, HBD/HBB), glucose and cations (SLC2A1, SLC2A4, SLC2A14) in both active (SLC4A1, SLC4A2, SLC2A1, SLC2A4, SLC16A1, SLC2A14) and symporter forms (SLC2A1, SLC2A4, SLC2A14).

The oxidative status of these cells is also fundamental for perform their role systemically. Here, we identified ABCC4, CYB5R3, CAT, MPO, MTHFD1, EPX, FDXR, HMOX2, PRDX2, PRDX1 and CYBRD1 with oxidoreductase activity altered in the groups under comparison, especially CAT, MPO, EPX, PRDX1 and PRDX2 that present peroxidase and antioxidant activity.

Table III.2: Significant Gene Ontology terms retrieved by ProteinOn from the datasets of differentially expressed proteins in mRBC samples.

Term Name	%Occurence	e-value	info content	Proteins
structural constituent of cytoskeleton	5,66	1,1E-16	12,8025	SORBS3, KRT14, KRT1, EPB41, SPTB, ANK1, EPB42, KRT2, ACTB
glucose transmembrane transporter activity	1,89	4,0E-06	12,4066	SLC2A1, SLC2A4, SLC2A14
oxygen transporter activity	2,52	9,2E-06	10,3004	HBD, HBG1, HBA1, HBD/HBB
antioxidant activity	3,14	7,9E-04	7,8124	CAT, MPO, EPX, PRDX2, PRDX1
secondary active transmembrane transporter activity	3,77	1,9E-02	6,25025	SLC4A1, SLC4A2, SLC2A1, SLC2A4, SLC16A1, SLC2A14
ATPase activity	3,14	7,4E-01	4,68242	ABCC4, ATP1A1, ATP2B4, ABCA13, ABCB6
electron carrier activity	2,52	10 E-01	3,71219	CYB5R3, FDXR, HMOX2, derp12
oxidoreductase activity	7,55	1,0E+00	2,57999	ABCC4, CYB5R3, CAT, MPO, MTHFD1, EPX, FDXR, HMOX2, PRDX2, PRDX1, CYBRD1, derp12

But normal and flexible RBC are necessary to efficiently flow and sneak through narrow capillaries to deliver oxygen to cells ensuring a proper microenvironment for a tissue to perform its function. As mentioned, cytoskeleton organization accounts for the most altered molecular function for which several proteins contribute: SORBS3, KRT14, KRT1, EPB41, SPTB, ANK1, EPB42, KRT2 and ACTB. This observation provides extra evidence that malformed and abnormal RBC can contribute to CF pathology.

Pathway analysis performed by IPA's software highlighted similar results. Among the most significant networks and functions retrieved, we found small molecule biochemistry (in close correlation with oxidant/antioxidant status), cellular assembly and organization (related to cytoskeleton rearrangements) and molecular transport. Interestingly, a noteworthy pathway yielded by IPA has NFkB as a central molecule (Fig. III.2). Previous results in serum [35,36] and lung epithelial cells [37] of CF patients corroborate this result identifying the pathway where proinflammatory NFkB participates as significant in the CF's phenotype.

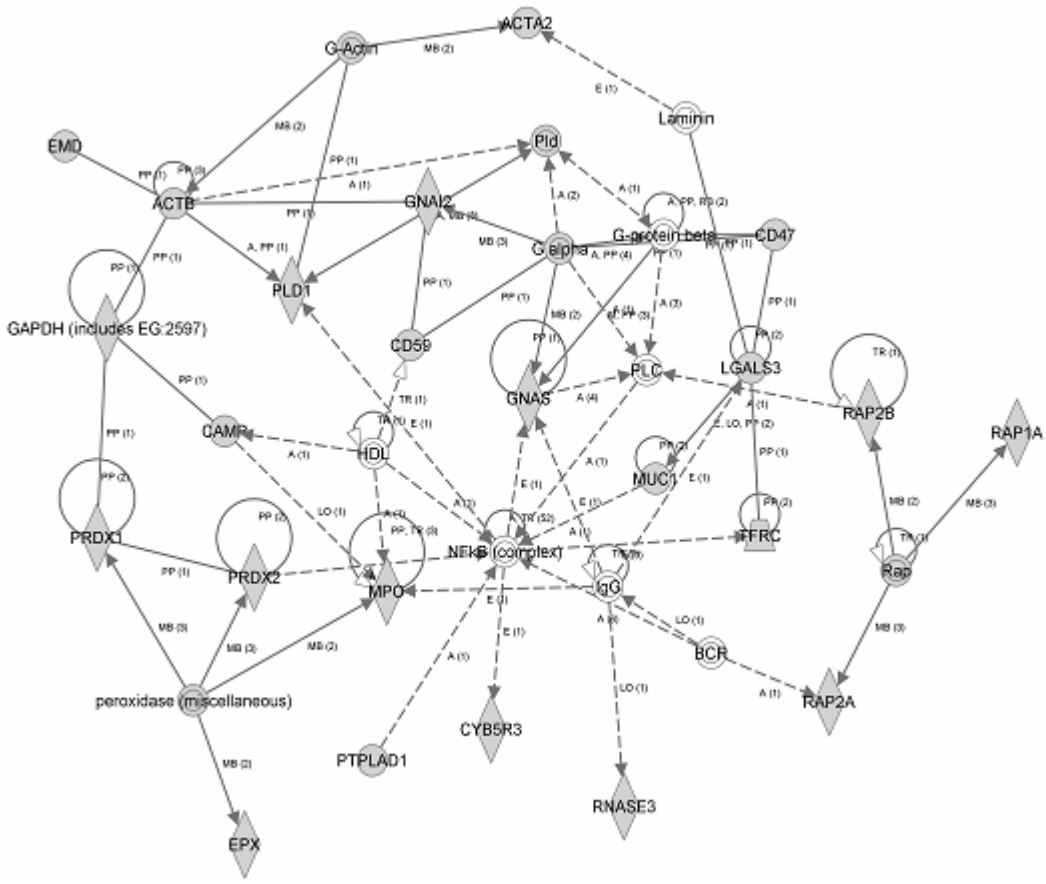


Figure III.2: IPA's molecular pathway where proinflammatoty NFkB occupies a central position among several proteins identified in mRBC. Several proteins related to antioxidant activity (MPO, EPX, PRDX1, PRDX2 and CAMP) are also emphasized. (in grey: proteins that were identified in the current work).

3.2.2 Tissue remodeling and antibacterial response in RBC

In CF, pulmonary parenchyma is constantly being challenged by persistent exposure to the action of proteolytic enzymes, such as neutrophilic elastase (ELANE), cathepsin G (CTSG) and proteinase 3 (PRTN3) that cause tissue injury characteristic of CF lung disease [38-40]. Besides being involved in tissue injury as member of the peptidase module of 22 neutrophil-derived proteins, CTSG inhibits IL-8, a known marker for inflammation in CF [41,42]. In our results, CTSG presents increased abundance in CF carriers when compared to controls but its levels significantly drop when comparing CF patients with CF carriers and healthy non-CF individuals. This effect might result in inefficient inhibition of IL-8-dependent proinflammatory pathway promoting the characteristic inflammatory phenotype evidenced by the patients.

We were able to identify differential regulation in several proteins involved in antibacterial response: Lysozyme C (LYZ), Cathelicidin antimicrobial peptide (CAMP), lactotransferrin (LTF) and CTSG, already described as participating severe acute respiratory syndorm (SARS) or pneumonitis. All these proteins belong to the peptidase module compromising 22 neutrophil-derived protein and have primarly a bacteriolytic function, actively involved in deplete the airways from opportunistic pathogens that colonize CF's lungs [43-46]. Both LTF and CTSG have already been associated with CF. Also, CTSG has antibacterial activity against the Gram-negative bacterium *P.aeruginosa* and inhibition by LPS from the same bacteria [47]. Curiously, LYZ, LTF and CAMP were only identified in the CF's carriers' group, each of them by 2 peptides only. As mentioned, our sample preparation resulted in an enrichment of proteins from the mRBC proteome. We cannot exclude however a small percentage of cross contamination with other blood cells, resulting therefore in the identification of such proteins with low peptides count. Moreover, these three proteins present a primarly antibacterial function. Lack of their identification at systemic level might indicate activity at other locations where infection takes place, such as the lung. We believe that using other type os biological sample, like one derived fro the respiratory system, will unravel the presence of these proteins in sites of infection.

3.2.3 Oxidative stress in mRBC

The erythrocytes are naturally and constantly exposed to endogenous oxidants such superoxide and hydrogen peroxide. To overcome the damage that it might create, antioxidant defenses are of great importance for maintenance of normal cellular functions. Catalase (CAT) is a homotetrameric antioxidant enzyme and serves to protect cells from the toxic effects of hydrogen peroxide converting this oxidant into water and oxygen. It occurs in almost all aerobically respiring organisms and highly concentrated and active in RBC. Besides its role in detoxifying hydrogen peroxide, CAT promotes growth of cells including T-cells, B-cells, myeloid leukemia cells, melanoma cells, mastocytoma cells, normal and transformed fibroblast cells with implication in hyperoxia and pulmonary defense, especially at the alveolar level [48]. CAT's activity is highly correlated with both superoxide dismutase-1 (SOD1) and glutathione peroxidase (GPX). Hydrogen peroxide is formed from the dismutation of two radical anion superoxide ($O_2^{\bullet-}$) catalised by SOD1 and its detoxification can be achieved by

catalase or GPX though *in vitro* studies demonstrate preferential catalase's pathway [49,50]. In our study, neither SOD1 nor GPX were identified, probably due to their hydrophilic character and therefore not enriched in our methodology, their low abundance or technical constrains, but we aim to determine their antioxidant's activity by a complementary biochemical method to better elucidate this. In a similar fashion, it was demonstrated that GPX presented lower levels in nasal epithelial cells of CF patients compared with healthy though severe patients present higher levels of the enzyme than mild ones suggesting that GPX might constitute a first-line defence against oxidative damage in earlier stages of the disease (for detailed results, see Chapter V).

We also identified peroxiredoxin 1 and 2 (PRDX 1 and PRDX 2, respectively) as differentially expressed in our study. This family of peroxidases are involved in redox regulation of the cell by eliminating peroxides by conjugation with reducing equivalents provided through the thioredoxin system. In the lungs, peroxiredoxins have been implicated in protection against exogenous as well as endogenous oxidant challenge [51] and particularly PRDX2 has been found elevated in lung carcinomas [52]. PRDX1, on the other hand, was found elevated in mRBC of sickle cell patients [5]. Additionally to their antioxidant properties, peroxiredoxins might participate in the signalling cascades of growth factors and TNF-α by regulating the intracellular concentrations of hydrogen peroxide. TNF-α is the main proinflammatory cytokine secreted by inflammatory macrophages and makes a strong contribution to the proinflammmatory phenotype of the CF airways [36].

CF's respiratory tract is colonized/infected at early stages by opportunistic pathogens resulting increased phagocytic responses and active secretion of reactive intermediates by neutrophilic myeloperoxidase (MPO) [53,54]. Due to highly infiltration of neutrophils in CF respiratory surfaces, large amounts of MPO are release constituting the predominant peroxidase at that site [55]. Although catalase can catalyze detoxification of hydrogen peroxide, MPO converts this compound into potent bactericidal oxidant hypochlorous acid (HOCl), indicating that this enzyme can perpetuate oxidative damage [56]. Moreover, MPO also acts as a phagocyte-derived nitric oxide (NO) oxidase that prevents its functions as a bronchodilator, bacterostatic agent and anti-inflammatory mediator [55]. Elevated levels of MPO have been reported in CF's BALF [57], sputum [56,58,59] and MPO was also supposed to be an autoantigen for the disease [59]. This NO oxidase is however not restricted to MPO.

Eosinophil peroxidase (EPX) is an antibacterial protein arising from exocytosed granules of activated eosinophils in response to hypersensitive responses, such as allergen-induced asthma [60], and parasite invasion [61,62]. Although eosinophils numbers do not seem to be elevated in CF [63,64], it is noteworthy to observe that heme peroxidase activity is several orders of magnitude lower in asthma when compared to levels reported in CF. Eosinophils-specific granule proteins are marked citotoxic for parasites but its oxidant products might interact with the surronding molecules causing tissue injury as demonstrated in pneumocytes [65,66].

3.2.4 Structural organization of mRBC

The ability of the RBC to perform their function is dependent of their plasticity and mechanical deformability to pass through blood vessels' walls. Cytoskeleton rearrangements are therefore crucial to guarantee the maintenance of an adequate morphology. In our study, several proteins related to such functions were identified with possible consequences for CF.

Flotilins (FLOT) are major integral proteins of erythrocyte lipid rafts [67] that participate in signal transduction and intracellular trafficking processes, such as caveolar-mediated endocytosis and cell adhesion [68]. Closely associated with these is stomatin (STOM), a detergent-resistant protein that associates with lipid rafts rather than binding to the cytoskeleton to act as a regulator of monovalent cation flux across membrane [69]. Work from Salzer and Prohaska showed that stomatin and the flotilins are the most abundant integral proteins of the RBC rafts that are independently organized in high-order oligomers, acting as a separate scaffold at their cytoplasmic face [67]. In our study, stomatin was the second protein identified by higher number of peptides (only overcome by band 3 anion transport protein) but its levels were similar across the groups under comparison. On the other hand, FLOT 1 and FLOT2 presented similar expression patterns: decrease in abundance in CF carriers and patients when compared with controls and a non-significant increase in patients with severe lung disease in comparison with mild ones.

Glycophorin A and C (GYPA and GYPC, respectively) are mRBC proteins also identified as differentially expressed here: GYPA presented gradually lower expression from healthy non-CF to CF carriers and from those to CF patients while levels of GYPC

were similar in controls and carriers and decreased abundance in CF patients. Both proteins were not significantly different between severe and mild patients. GYPA is the major intrinsic membrane protein of the erythrocyte influencing function, high activity and translocation of band 3 anion transport protein (SLC4A1) while GYPC is a minor sialoglycoprotein in human RBC with important role in regulating their stability. All STOM, GYPA, GYPC and SLC4A1 associate with spectrin-actin-based cytoskeleton via linker proteins.

The mRBC is an organized structure that results from the direct interaction of a spectrin-actin-based membrane skeletal network with the lipid bilayer or by association with linker proteins that interact simultaneously with the cytoplasmic domain of transmembrane proteins and spectrin [70]. Cytoskeleton spectrins α and β (STPA1 and SPTB, respectively) were identified here as less abundant in CF carriers and patients, with similar levels in these two groups, while actin was also less abundant in CF carriers and patients when compared with non-CF. These proteins associate with band 4.1 to form the cytoskeletal superstructure of the mRBC. Linker proteins ankyrin and erythrocyte membrane protein band 4.1 and 4.2 were identified at similar levels in CF and non-CF individuals with lower expression at the CF carriers' group while β-adducin present lower abundance in CF carriers and patients (with similar levels in these two groups) than in healthy non-CF. Linker proteins are important in regulating membrane physical properties of mechanical stability and deformability by promoting and stabilizing spectrin-actin interaction. Overall, a tendency of lower abundance of proteins from the cytoskeleton, the spectrin-actin network and linker proteins is observed in CF patients when compared with the other groups. Taking into account all these differences in abundances of proteins from the cytoskeleton and spectrins' network, we postulate an imbalanced and impaired organization of the RBC scaffold with consequences in terms of RBC deformability and ability to efficiently perform their function. Lack of functional and malleable RBC is in the basis of reduced deformation-induced ATP release and consequent increase in pulmonary hypertension following deterioration of pulmonary function.

3.2.5 Differential identification of hemoglobins' subunits

During the isolation of mRBC proteins, incubation in basic medium, washes and solubilization in a methanol-buffered solution are performed to eliminate evident contamination of the microssomal fraction with hemoglobin. However, and although total removal of this protein is virtually impossible, its appearance in the MS spectra can actually be related with differential abundance in the samples under analysis.

It is known that hemoglobin has more functions than transport oxygen from the lung to the various peripheral tissues receiving carbon dioxide in exchange. Several published studies have highlighted the role of hemoglobin in pathological conditions [71-74]. Moreover, in a recent study from our group using a SELDI-TOF approach combined with protein's purification and identification by MALDI-TOF/TOF MS, hemoglobin was identified as differentially expressed between CF patients, controls and asthmatics [75]. Since we had extra care in excluding samples with evident hemolysis, we believe that this difference is biological. Here, three subunits of hemoglobin were identified more abundantly in CF carriers and CF patients: hemoglobin subunit α, β and δ (HBA1, HBB and HBD, respectively). Patients with severe lung disease also presented higher levels of these proteins than mild ones. Oxidative stress is one of the imbalanced processes in CF and previous works have associated higher oxidative stress with higher rate of hemolysis [76]. The identification of higher levels of hemoglobin's subunits in mRBC of CF patients is in accordance with previous results [75] and might reveal an increased susceptibility of CF's RBC to oxidative damage [76,77].

4. SUMMARY

By using a high throughput MS-based approach we identified here several differentially expressed proteins in CF´s mRBC associated with impaired antioxidant defence, increased lipid peroxidation and mechanical/deformability processes, confirming inability of CF's mRBC to efficiently respond against aggressions nor to mechanical deformation and promote vessel dilation preventing pulmonary hypertension. We presented here, for the first time, CFTR identification in mRBC by a MS-based approach. However, this identification was only possible in mRBC from patients homozygous for ΔF508 probably due to intracellular accumulation of this misfolded protein by the cell quality control.

REFERENCES

[1] Goodman SR, Krebs KE, Whitfield CF, Riederer BM, Zagon IS. Spectrin and related molecules. CRC Crit Rev Biochem 1988; 23:171-234.

[2] Rabilloud T, Blisnick T, Heller M, Luche S, Aebersold R, *et al.* Analysis of membrane proteins by two-dimensional electrophoresis: comparison of the proteins extracted from normal or Plasmodium falciparum-infected erythrocyte ghosts. Electrophoresis 1999; 20:3603-10.

[3] Low TY, Seow TK, Chung MC. Separation of human erythrocyte membrane associated proteins with one-dimensional and two-dimensional gel electrophoresis followed by identification with matrix-assisted laser desorption/ionization-time of flight mass spectrometry. Proteomics 2002; 2:1229-39.

[4] Kakhniashvili DG, Bulla LA, Jr., Goodman SR. The human erythrocyte proteome: analysis by ion trap mass spectrometry. Mol Cell Proteomics 2004; 3:501-9.

[5] Kakhniashvili DG, Griko NB, Bulla LA, Jr., Goodman SR. The proteomics of sickle cell disease: profiling of erythrocyte membrane proteins by 2D-DIGE and tandem mass spectrometry. Exp Biol Med (Maywood) 2005; 230:787-92.

[6] Tyan YC, Jong SB, Liao JD, Liao PC, Yang MH, *et al.* Proteomic profiling of erythrocyte proteins by proteolytic digestion chip and identification using two-dimensional electrospray ionization tandem mass spectrometry. J Proteome Res 2005; 4:748-57.

[7] Pasini EM, Kirkegaard M, Mortensen P, Lutz HU, Thomas AW, *et al.* In-depth analysis of the membrane and cytosolic proteome of red blood cells. Blood 2006; 108:791-801.

[8] D'Alessandro A, Righetti PG, Zolla L. The red blood cell proteome and interactome: an update. J Proteome Res 2010; 9:144-63.

[9] Cystic Fibrosis Foundation. Cystic Fibrosis Foundation, Patient Registry 2005 Annual Report, Bethesda, MD. 2004.

[10] Welsh MJ, Smith AE. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. Cell 1993; 73:1251-4.

[11] Dominguez C, Gartner S, Linan S, Cobos N, Moreno A. Enhanced oxidative damage in cystic fibrosis patients. Biofactors 1998; 8:149-53.

[12] Collins CE, Quaggiotto P, Wood L, O'Loughlin EV, Henry RL, *et al.* Elevated plasma levels of F2 alpha isoprostane in cystic fibrosis. Lipids 1999; 34:551-6.

[13] Sprague RS, Ellsworth ML, Stephenson AH, Kleinhenz ME, Lonigro AJ. Deformation-induced ATP release from red blood cells requires CFTR activity. Am J Physiol 1998; 275:H1726-32.

[14] Sterling KM, Jr., Shah S, Kim RJ, Johnston NI, Salikhova AY, *et al.* Cystic fibrosis transmembrane conductance regulator in human and mouse red blood cell membranes and its interaction with ecto-apyrase. J Cell Biochem 2004; 91:1174-82.

[15] Verloo P, Kocken CH, Van der Wel A, Tilly BC, Hogema BM, *et al.* Plasmodium falciparum-activated chloride channels are defective in erythrocytes from cystic fibrosis patients. J Biol Chem 2004; 279:10316-22.

[16] Stumpf A, Wenners-Epping K, Walte M, Lange T, Koch HG, *et al.* Physiological concept for a blood based CFTR test. Cell Physiol Biochem 2006; 17:29-36.

[17] Lange T, Jungmann P, Haberle J, Falk S, Duebbers A, *et al.* Reduced number of CFTR molecules in erythrocyte plasma membrane of cystic fibrosis patients. Mol Membr Biol 2006; 23:317-23.

[18] Lader AS, Prat AG, Jackson GR, Jr., Chervinsky KL, Lapey A*, et al.* Increased circulating levels of plasma ATP in cystic fibrosis patients. Clin Physiol 2000; 20:348-53.
[19] Sprague RS, Ellsworth ML, Stephenson AH, Lonigro AJ. ATP: the red blood cell link to NO and local control of the pulmonary circulation. Am J Physiol 1996; 271:H2717-22.
[20] Sprague RS, Olearczyk JJ, Spence DM, Stephenson AH, Sprung RW*, et al.* Extracellular ATP signaling in the rabbit lung: erythrocytes as determinants of vascular resistance. Am J Physiol Heart Circ Physiol 2003; 285:H693-700.
[21] Lundberg JO, Nordvall SL, Weitzberg E, Kollberg H, Alving K. Exhaled nitric oxide in paediatric asthma and cystic fibrosis. Arch Dis Child 1996; 75:323-6.
[22] Liang G, Stephenson AH, Lonigro AJ, Sprague RS. Erythrocytes of humans with cystic fibrosis fail to stimulate nitric oxide synthesis in isolated rabbit lungs. Am J Physiol Heart Circ Physiol 2005; 288:H1580-5.
[23] Johnson C, Butler SM, Konstan MW, Morgan W, Wohl ME. Factors influencing outcomes in cystic fibrosis: a center-based analysis. Chest 2003; 123:20-7.
[24] Pacheco P, Detection rate of CF causing CFTR mutations in Portugal. in: WHO, (Ed.), The Molecular Genetic Epidemiology of Cystic Fibrosis, Human Genetics Programme, Chronic Diseases and Health Promotion, 2004.
[25] Dodge JT, Mitchell C, Hanahan DJ. The preparation and chemical characteristics of hemoglobin-free ghosts of human erythrocytes. Arch Biochem Biophys 1963; 100:119-30.
[26] Blonder J, Chan KC, Issaq HJ, Veenstra TD. Identification of membrane proteins from mammalian cell/tissue using methanol-facilitated solubilization and tryptic digestion coupled with 2D-LC-MS/MS. Nat Protoc 2006; 1:2784-90.
[27] Rudnick PA, Clauser KR, Kilpatrick LE, Tchekhovskoi DV, Neta P*, et al.* Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses. Mol Cell Proteomics 2010; 9:225-41.
[28] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982; 157:105-32.
[29] Sonnhammer EL, von Heijne G, Krogh A. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol 1998; 6:175-82.
[30] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001; 305:567-80.
[31] Medina-Aunon JA, Paradela A, Macht M, Thiele H, Corthals G*, et al.* Protein Information and Knowledge Extractor: Discovering biological information from proteomics data. Proteomics 2010; 10:3262-71.
[32] Faria D, Pesquita C, Couto F, Falcão A, ProteinOn: Web Tool for protein Semantic Similarity, ISMB/ECCB, 2007.
[33] Chang XB, Hou YX, Jensen TJ, Riordan JR. Mapping of cystic fibrosis transmembrane conductance regulator membrane topology by glycosylation site insertion. J Biol Chem 1994; 269:18572-5.
[34] Gomes-Alves P, Couto F, Pesquita C, Coelho AV, Penque D. Rescue of F508del-CFTR by RXR motif inactivation triggers proteome modulation associated with the unfolded protein response. Biochim Biophys Acta 2010; 1804:856-865.
[35] Charro N, Hood BL, Faria D, Pacheco P, Azevedo P*, et al.* Serum proteomics signature of Cystic Fibrosis patients: A complementary 2-DE and LC-MS/MS approach. J Proteomics 2011.

[36] Srivastava M, Eidelman O, Jozwik C, Paweletz C, Huang W, *et al.* Serum proteomic signature for cystic fibrosis using an antibody microarray platform. Mol Genet Metab 2006; 87:303-10.
[37] Tabary O, Boncoeur E, de Martin R, Pepperkok R, Clement A, *et al.* Calcium-dependent regulation of NF-(kappa)B activation in cystic fibrosis airway epithelial cells. Cell Signal 2006; 18:652-60.
[38] Delacourt C, Le Bourgeois M, D'Ortho MP, Doit C, Scheinmann P, *et al.* Imbalance between 95 kDa type IV collagenase and tissue inhibitor of metalloproteinases in sputum of patients with cystic fibrosis. Am J Respir Crit Care Med 1995; 152:765-74.
[39] Witko-Sarsat V, Delacourt C, Rabier D, Bardet J, Nguyen AT, *et al.* Neutrophil-derived long-lived oxidants in cystic fibrosis sputum. Am J Respir Crit Care Med 1995; 152:1910-6.
[40] Vandivier RW, Fadok VA, Hoffmann PR, Bratton DL, Penvari C, *et al.* Elastase-mediated phosphatidylserine receptor cleavage impairs apoptotic cell clearance in cystic fibrosis and bronchiectasis. J Clin Invest 2002; 109:661-70.
[41] Ordonez CL, Henig NR, Mayer-Hamblett N, Accurso FJ, Burns JL, *et al.* Inflammatory and microbiologic markers in induced sputum after intravenous antibiotics in cystic fibrosis. Am J Respir Crit Care Med 2003; 168:1471-5.
[42] Sagel SD, Accurso FJ. Monitoring inflammation in CF. Cytokines. Clin Rev Allergy Immunol 2002; 23:41-57.
[43] Gibson RL, Emerson J, McNamara S, Burns JL, Rosenfeld M, *et al.* Significant microbiological effect of inhaled tobramycin in young children with cystic fibrosis. Am J Respir Crit Care Med 2003; 167:841-9.
[44] Ramsey BW. Management of pulmonary disease in patients with cystic fibrosis. N Engl J Med 1996; 335:179-88.
[45] Baltimore RS, Christie CD, Smith GJ. Immunohistopathologic localization of Pseudomonas aeruginosa in lungs from patients with cystic fibrosis. Implications for the pathogenesis of progressive lung deterioration. Am Rev Respir Dis 1989; 140:1650-61.
[46] Thomson E, Brennan S, Senthilmohan R, Gangell CL, Chapman AL, *et al.* Identifying Peroxidases and Their Oxidants in the Early Pathology of Cystic Fibrosis. Free Radic Biol Med 2010.
[47] Wasiluk KR, Skubitz KM, Gray BH. Comparison of granule proteins from human polymorphonuclear leukocytes which are bactericidal toward Pseudomonas aeruginosa. Infect Immun 1991; 59:4193-200.
[48] Rahman I, Biswas SK, Kode A. Oxidant and antioxidant balance in the airways and airway diseases. Eur J Pharmacol 2006; 533:222-39.
[49] Chang P, Cheng E, Brooke S, Sapolsky R. Marked differences in the efficacy of post-insult gene therapy with catalase versus glutathione peroxidase. Brain Res 2005; 1063:27-31.
[50] Ray G, Husain SA. Oxidants, antioxidants and carcinogenesis. Indian J Exp Biol 2002; 40:1213-32.
[51] Lehtonen ST, Markkanen PM, Peltoniemi M, Kang SW, Kinnula VL. Variable overoxidation of peroxiredoxins in human lung cells in severe oxidative stress. Am J Physiol Lung Cell Mol Physiol 2005; 288:L997-1001.
[52] Lehtonen ST, Svensk AM, Soini Y, Paakko P, Hirvikoski P, *et al.* Peroxiredoxins, a novel protein family in lung cancer. Int J Cancer 2004; 111:514-21.

[53] Mohammed JR, Mohammed BS, Pawluk LJ, Bucci DM, Baker NR, *et al.* Purification and cytotoxic potential of myeloperoxidase in cystic fibrosis sputum. J Lab Clin Med 1988; 112:711-20.
[54] Van Der Vliet A, Nguyen MN, Shigenaga MK, Eiserich JP, Marelich GP, *et al.* Myeloperoxidase and protein oxidation in cystic fibrosis. Am J Physiol Lung Cell Mol Physiol 2000; 279:L537-46.
[55] Chapman AL, Morrissey BM, Vasu VT, Juarez MM, Houghton JS, *et al.* Myeloperoxidase-dependent oxidative metabolism of nitric oxide in the cystic fibrosis airway. J Cyst Fibros 2010; 9:84-92.
[56] Kettle AJ, Chan T, Osberg I, Senthilmohan R, Chapman AL, *et al.* Myeloperoxidase and protein oxidation in the airways of young children with cystic fibrosis. Am J Respir Crit Care Med 2004; 170:1317-23.
[57] Gharib SA, Vaisar T, Aitken ML, Park DR, Heinecke JW, *et al.* Mapping the lung proteome in cystic fibrosis. J Proteome Res 2009; 8:3020-8.
[58] Worlitzsch D, Herberth G, Ulrich M, Doring G. Catalase, myeloperoxidase and hydrogen peroxide in cystic fibrosis. Eur Respir J 1998; 11:377-83.
[59] Sloane AJ, Lindner RA, Prasad SS, Sebastian LT, Pedersen SK, *et al.* Proteomic analysis of sputum from adults and children with cystic fibrosis and from control subjects. Am J Respir Crit Care Med 2005; 172:1416-26.
[60] Bousquet J, Chanez P, Lacoste JY, Barneon G, Ghavanian N, *et al.* Eosinophilic inflammation in asthma. N Engl J Med 1990; 323:1033-9.
[61] Gleich GJ. Mechanisms of eosinophil-associated inflammation. J Allergy Clin Immunol 2000; 105:651-63.
[62] Jong EC, Klebanoff SJ. Eosinophil-mediated mammalian tumor cell cytotoxicity: role of the peroxidase system. J Immunol 1980; 124:1949-53.
[63] Konstan MW, Hilliard KA, Norvell TM, Berger M. Bronchoalveolar lavage findings in cystic fibrosis patients with stable, clinically mild lung disease suggest ongoing infection and inflammation. Am J Respir Crit Care Med 1994; 150:448-54.
[64] Danel C, Erzurum SC, McElvaney NG, Crystal RG. Quantitative assessment of the epithelial and inflammatory cell populations in large airways of normals and individuals with cystic fibrosis. Am J Respir Crit Care Med 1996; 153:362-8.
[65] Agosti JM, Altman LC, Ayars GH, Loegering DA, Gleich GJ, *et al.* The injurious effect of eosinophil peroxidase, hydrogen peroxide, and halides on pneumocytes in vitro. J Allergy Clin Immunol 1987; 79:496-504.
[66] McElhinney B, Poynter ME, Shrivastava P, Hazen SL, Janssen-Heininger YM. Eosinophil peroxidase catalyzes JNK-mediated membrane blebbing in a Rho kinase-dependent manner. J Leukoc Biol 2003; 74:897-907.
[67] Salzer U, Prohaska R. Stomatin, flotillin-1, and flotillin-2 are major integral proteins of erythrocyte lipid rafts. Blood 2001; 97:1141-3.
[68] Hansen CG, Nichols BJ. Molecular mechanisms of clathrin-independent endocytosis. J Cell Sci 2009; 122:1713-21.
[69] Green JB, Fricke B, Chetty MC, von During M, Preston GF, *et al.* Eukaryotic and prokaryotic stomatins: the proteolytic link. Blood Cells Mol Dis 2004; 32:411-22.
[70] Pasini EM, Lutz HU, Mann M, Thomas AW. Red blood cell (RBC) membrane proteomics--Part I: Proteomics and RBC physiology. J Proteomics 2010; 73:403-20.

[71] Kozak KR, Su F, Whitelegge JP, Faull K, Reddy S*, et al.* Characterization of serum biomarkers for detection of early stage ovarian cancer. Proteomics 2005; 5:4589-96.
[72] Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. Lancet 2010; 376:2018-31.
[73] Vij R, Machado RF. Pulmonary complications of hemoglobinopathies. Chest 2010; 138:973-83.
[74] Owens EO. Endogenous carbon monoxide production in disease. Clin Biochem 2010; 43:1183-8.
[75] Gomes-Alves P, Imrie M, Gray RD, Nogueira P, Ciordia S*, et al.* SELDI-TOF biomarker signatures for cystic fibrosis, asthma and chronic obstructive pulmonary disease. Clin Biochem 2010; 43:168-77.
[76] Benabdeslam H, Abidi H, Garcia I, Bellon G, Gilly R*, et al.* Lipid peroxidation and antioxidant defenses in cystic fibrosis patients. Clin Chem Lab Med 1999; 37:511-6.
[77] Laskowska-Klita T, Chelchowska M. Antioxidant status in erythrocytes of cystic fibrosis children. Acta Biochim Pol 2001; 48:283-5.

Chapter IV

Proteomics of Human Nasal Epithelial Cells: a Molecular Portrait

Submitted with minor alterations to the Special Issue of Journal of Proteomics dedicated to the 4th EUPA Congress: Proteomics Odyssey Towards Next Decades.

Simões T[1,2,*], Charro N[1,2,*], Faria D[5], Pacheco P[3], Lopes C[4], Azevedo P[4], Bugalho de Almeida A[4], Couto FM[5], Chan KC[2], Isaaq HJ[2], Waybright T[2], Veenstra TD[2], Blonder J[2], Penque D[1]

[1]*Laboratory of Proteomics, Dept. of Genetics, INSA Dr. Ricardo Jorge, Lisbon, Portugal*
[2]*Laboratory of Proteomics and Analytical Technologies, SAIC-Inc., NCI at Frederick, MD, USA.*
[3]*Laboratory of Molecular Biology, Dept. of Genetics, INSA Dr. Ricardo Jorge, Lisbon, Portugal.*
[4]*Clínica Universitária de Pneumologia, Hospital de Santa Maria, Faculdade de Medicina da Universidade de Lisboa.*
[5]*Dept. of Informatics, Faculdade de Ciências da Universidade de Lisboa, Portugal.*

* These authors contributed equally to this work.

ABSTRACT

Here, a comprehensive proteomic investigation on nasal epithelial cells (NEC) using an improved subcellular fractionation method to obtain soluble- and membrane-enriched subproteomes followed by a high sensitive multidimensional protein identification technology (MudPIT) is described. NEC cells were collected by a non-invasive brushing procedure from 129 healthy subjects [84F, 45M; (36±9) years]. Soluble- and membrane-enriched protein fractions were isolated by differential centrifugation and methanol-based extraction. Tryptic digests were fractionated by strong cation-exchange (SCX) liquid chromatography (LC) followed to microcapillary reversed phase (RP) LC coupled directly to tandem mass spectrometer (MS/MS) analysis.

In a set of 1482 identified proteins, 947 (63.9%) proteins were annotated as membrane associated proteins according to Gene Ontology (GO). The grand average hydropath value index (GRAVY), the transmembrane mapping predictions and Human Protein Reference Database (HPRD) information confirmed an enrichment of hydrophobic proteins in membrane subproteome. Exponentially modified Protein Abundance Index (emPAI) was used to estimate total protein content in NEC highlighting mitochondrial, ribosomal and cytoskeleton proteins in the most abundant proteins' group and immunological, molecular transport, morphology or cellular assembly and organization in the lower abundant proteins' group. Ingenuity pathway analysis revealed an enrichment of molecular and cellular functions associated with cell death, protein folding and drug metabolism in soluble fraction while in membrane fraction there was a predominance in functions associated with molecular transport, protein trafficking and cell-to-cell signalling and interaction. Comparing our data with those described in the literature for other respiratory tract epithelium revealed some molecular similarities that

include proteins associated to maintenance of physical barrier and immunological defence, providing extra support to the use of NEC in the assessment of outcome and treatment response of airway diseases such as asthma, cystic fibrosis and chronic obstructive pulmonary disease.

Keywords:
nasal epithelium; upper and lower airways; cellular fractionation; two-dimensional liquid chromatography; multidimensional protein identification technology (MudPIT), mass spectrometry; functional analysis

BACKGROUND

Nasal and bronchial mucosae consist of a pseudostratified ciliated columnar epithelium resting on a basement membrane. Beneath the basement membrane is the submucosa, containing blood vessels, mucous glands, structural cells, nerves and some inflammatory cells [1]. Its primarily function is to humidify, heat or cool and clean inhaled air before reaching alveolar compartment. It is controversial the assessment of nasal epithelium (NE) as surrogate model of lower airways' characteristics and functions [2]. Some authors say that the NE is completely unrepresentative of lungs epithelial layer due to differences in embryonic origins [3], anatomy [4], cellular populations [5-8] and transepithelial resistances [9]. The main difference between the upper and lower airways is that upper airway patency is largely influenced by vascular tone, whereas, in the lower airway, airflow is influenced predominantly by smooth muscle contraction. Nevertheless, both portions act as a transport system moving air in and out of the lungs. However, there seem to be more similarities than differences between upper and lower airways that have been extensively reported [10-14]. In fact, the concept of united airways is being increasingly used highlighting the fact that there is an anatomic and physiological continuity between the upper and lower airways [10] where reactions in the first influence complications in the latter [4,15].

It is unquestionable the importance of airway epithelium as a physical barrier between external environment and internal milieu [16,17]. Under physiological conditions, airway epithelium, in coordination with airway surface liquid (ASF) and mucous, forms a highly regulated physical ad functional barrier between these two environments, sensing and actively responding to alterations in its surroundings therefore controlling many responses typical of respiratory diseases [3].

Frequently exposed to external challenges/aggression, respiratory epithelium has developed various defence mechanisms to protect itself and thus the organism from infections and inhaled particles. These defence mechanisms depend on mucociliary clearance via airway ciliated cells, on the capacity to regenerate an intact respiratory epithelium after injury and its ability to participate in inflammatory responses by secreting pro- and anti-inflammatory cytokines, among others [18]. In upper and lower airways, responses are triggered by similar mechanisms including irritants (smoke, cold air or exercise) and allergens (pollens or mold) leading to narrowing of the lumen, mucous secretion and comparable patterns of inflammation that depend on the circadian rhythm.

Respiratory epithelium is thus crucial to maintain the normal respiratory function and any alteration of its structure and function may play an important role in the pathogenesis of the chronic lung diseases [19].

In the monogenetic disease Cystic Fibrosis (CF), patients carry mutations in the CF transmembrane conductance regulator (CFTR) gene that codes for a defective protein [20,21] which jeopardized the functions of epithelia in several organs [22] specially at lung's level [23,24]. On the other hand, asthma displays significant heterogeneity and variability in its clinical expression due to complex and interactive factors. Environmental factors underlying asthma inception can range from viral respiratory tract infections in infancy to occupation exposures in adults [25]. According to the current understanding, many of the persistent inflammatory and structural responses in patients with asthma, including airway allergen sensitization in genetically susceptible individuals, could follow from a defective epithelium, leading to a chronic wound response and repeated environmental injury [18]. Thus, it seems that airway remodelling, especially at the level of airways epithelia [6,13], is common to both diseases implying a dynamic process of repair after injury/inflammation associated with deposit and degradation of extracelular matrix.

To our knowledge, only two studies aimed to investigate the NE by proteomics [26,27], one of them by our own group. These investigations were 2DE-based and identification was achieved by using MALDI-TOF MS with around 60 protein spots being excised and identified.

Herein, we report the use of subcellular fractionation combined with improved methods of solubilization and digestion (namely methanol-based buffer for solubilization and tryptic digestion of membrane proteins), followed by SCX-RPLC coupled with MS/MS for large-scale profiling of NE soluble- and membrane-enriched proteome as a complementary or alternative approach to 2DE-MALDI-TOF MS.

Overall, we identified 1482 different proteins: 313 and 947 proteins in soluble and membrane fraction, respectively, with an overlap of 222 proteins. 43 proteins were classified as uncharacterized or *in-silico* predicted only. Bioinformatic analysis correctly allocated the identified proteins into soluble or membrane compartments thus confirming an enrichment of proteins in the respective subfraction and validating the efficacy of the methodology applied. Proteins were annotated according to their primary localization, biological processes and functions highlighting important roles of epithelium.

Given the well-known similarities in mucosal responses between upper and intrathoracic airways described before, proteomic profiles observed in the nose may also provide insights into the pathology of airway diseases [28] such as Asthma, Chronic Obstructive Pulmonary Disease (COPD) and Cystic Fibrosis (CF).

2. MATERIALS AND METHODS

2.1 Demographics and Sample Collection

Ethical approval and informed consent were obtained from participating institutions and enrolled individuals, respectively. In total, 129 volunteers [84 F, 45 M; (36±9) years] clinically characterized were recruited for sample collection. Nasal cells were harvested by a non-invasive brushing procedure previously described [29]. Cells were recovered immediately after brushing and washed with ice-cold PBS. An aliquot was saved at 4ºC for cellular integrity analysis and the remaining was pelleted by centrifugation at 3000 rpm for 5 min at ambient temperature recovering around 1-2x10^6 cells and preserved at -80ºC until analysis.

2.2 Cellular Integrity Analysis

Cellular integrity was evaluated by cell fixing in 3.7% (v/v) formaldehyde, 3.7% (w/v) sucrose in PBS, centrifuged for 5 min at 1000 rpm in a CytoSpin 3 centrifuge (Shandon; Life Sciences International, Cheshire, UK) to adhere to silane-coated glass slides (Sigma, St Louis, Missouri) and staining according to Leishman method [30] using freshly prepared reagent. Cellular preparations were qualitatively evaluated by conventional light microscopy to guarantee enough intact cells for further analysis while avoiding evident contamination with blood cells from occasional bleeding during sample collection.

2.3 Sample fractionation by ultracentrifugation

Briefly, pelleted nasal cells were thawed on ice, pooled together and resuspended in 1mL of lysis buffer [10mM Tris-HCl (pH 7.6), 1mM EDTA and protease inhibitors cocktail (P8340-5mL, Sigma)]. Cell lysis was assisted by

intermittent sonication cycles (10x10'' pulses interleaved by 30'') on ice and sample was clarified by centrifugation at 2000x*g* for 3 min at 4ºC to remove cellular debris.

Proteins from the soluble fraction were obtained from this clarified sample by ultracentrifugation at 100 000x*g* for 90 min at 4ºC. The obtained pellet was resuspended in 50mM NH_4HCO_3 and sonicated before another ultracentrifugation in the previously mentioned conditions. Supernatant was discarded and pellets homogenized and incubated in 100mM Na_2CO_3 for 2h at 4ºC on a rotisserie followed by similar ultracentrifugation conditions. Resulting pellet was washed twice with ddH_2O to obtain the membrane-enriched proteins. Protein concentrations were established by BCA Protein Assay Kit (Pierce). 100 µg of lyophilized protein from soluble or membrane fraction were resuspended in either 50mM NH_4HCO_3 or 50mM NH_4HCO_3/60% (v/v) methanol [31], respectively, and sequence-grade modified trypsin (Promega, Madison, WI, USA) was added at an enzyme:protein ratio of 1:50 to soluble proteins' fraction and 1:20 to membrane proteins' fraction for overnight in-solution proteolysis. Tryptic peptides were desalted by SPE (3M™ Empore™ High Performance Extraction Disk Cartridges) according to manufacturer's instructions and lyophilized to dryness prior to 2D-LC-MS/MS analysis.

2.4 SCX and RP Fractionation

The first separation involved off-line peptide fractionation by strong cation exchange (SCX) liquid chromatography (LC) using in-house packed nanocolumns as described [31]. Desalted lyophilized peptides were solubilized in 45% (v/v) ACN/0.1% (v/v) FA to achieve a concentration of approximately 0.5-1 µg per µL. Peptides were eluted with an NH_4HCO_2 multistep gradient at a flow rate of 200 µL/min as follows: 0-1% B in 2 minutes, 1-10% B in 60 minutes, 10-62 % B in 20 minutes, 62-100% B in 3 minutes (Mobile phase A: 45% (v/v) ACN; Mobile phase B: 45%(v/v) ACN/0.5M NH_4HCO_2, pH 3). Elution monitoring was achieved by native fluorescence with excitation wavelength of 266nm at an emission wavelength of 340nm. In total, 96 fractions were obtained for each SCX separation and pooled to final 10 fractions intelligently assembled based on peptides' peak distributions and intensities. The SCX-LC fractions were lyophilized to dryness and reconstituted in 0.1% (v/v) FA immediately prior to reverse phase separation. Each SCX fraction was analysed in

duplicate by microcapillary reverse phase chromatography coupled online to a linear ion trap mass spectrometer (LTQ, ThermoElectron, ThermoFisher Scientific Inc., San Jose, CA) using the nanoelectrospray ionization source supplied by the manufacturer. After injecting 5 µL of sample, the column was washed for 30 minutes (at 0.5 µL/min) with 2% B and peptides eluted (at 0.25 µL/min) using a linear gradient as follows: 2-60% B in 100 minutes, 60-98% B in 20 minutes, 98% B for 20 minutes. The column was re-equilibrated with 2% B for 30 minutes prior to subsequent sample loading using the flow rate of 0.5 µL/min. (Mobile phase A: 0.1% (v/v) FA; Mobile phase B: 0.1% (v/v) FA in ACN. The LIT mass spectrometer was operated in a data dependent mode in which each full MS scan (scan range *m/z* 350-1800) was followed by five MS/MS scans, where the five most abundant peptide molecular ions detected from the MS scan were dynamically selected for five subsequent MS/MS scans using a normalized collision-induced dissociation (CID) energy of 35% and a dynamic exclusion of 60secs to reduce redundant selection of peptides.

2.5 Data Processing and Bioinformatic Analysis

The CID spectra were analyzed using SEQUEST^{TM} operating on a Beowulf 18-node parallel virtual machine cluster computer (ThermoElectron, Thermo Fisher Scientific, Waltham, MA, USA) using a UniProt non-redundant human proteome database (http://www.expasy.org, 03/2008 release). Only peptides with conventional tryptic termini (allowing for up to two internal missed cleavages) possessing delta-correlation scores (ΔCn) >0.1 and charge state-dependent cross correlation (Xcorr) criteria of >2.1 for $[M+H]^{1+}$, >2.3 for $[M+H]^{2+}$ and >3.5 for $[M+H]^{3+}$ were considered as legitimate identifications. Results were further analysed using in-house developed software for determination of unique peptides and proteins, considering only positive identifications when at least 2 unique peptides (UPs) *per* protein were assigned. The obtained identifications are the sum of 70 and 80 SCX fractions for the soluble and membrane-enriched fractions, respectively, each analysed in duplicate by RP. A final list of the identified proteins with the corresponding number of unique (non-redundant) peptides (UPs) and total count for those peptides (TPs) observed in two consecutive runs was obtained. To increase identifications' accuracy and confidence, only proteins identified in three independent experiments were positively identified and representative

of NEC. The false discovery rate (FDR) at peptide level was calculated by searching the data against a decoy database and was estimated to be between 4.2 and 4.5%.

To aid the biological interpretation of the extensive proteins' lists generated in these studies, proteins were categorized using the Gene Ontology, curated information deposited in Human Protein Reference Database (http://www.hprd.org, release 8, July 2009) and Ingenuity Pathway Analysis (Ingenuity® Systems, www.ingenuity.com) knowledgebase to retrieve information on biological processes, molecular functions and primary location. The mapping of putative transmembrane domains was performed in Centre for Biological Sequence TMHMM server (http://www.cbs.dtu.dk/services/TMHMM) and the grand average hidropathy values (GRAVY index) [32] were calculated using Gene Infinity tools (http://www.geneinfinity.org/sms_proteingravy.html).

Absolute protein content in the samples was estimated by exponentially modified protein abundance index (emPAI) [33], an index that compares the number of parent ions *per* protein observed in one's experiments with the number of predicted peptides that the correspondent protein origins. emPAI normalization between experiments was calculated to determine protein contents in molar percentages, as described by Ishihama *et al.* 2005 [33].

3. RESULTS AND DISCUSSION

3.1 General characterization of the identified proteins in NEC's subproteomes

Herein, we report results of a proteome-wider characterization of the NEC proteome using a high-throughput strategy that relies on cell fractionation coupled with SCX-RPLC-MS/MS that may facilitate discovery of respiratory biomarkers and/or therapeutic target discovery for respiratory diseases.

NEC specimens were obtained by a non-invasive nasal brushing procedure and their quality evaluated by cytology before used. The cytological analysis confirmed previous findings [29] by showing that most of recovered cells (about 85%) are epithelial, namely columnar, goblet and basal cells (results not shown).

Cellular fractionation into soluble and membrane-enriched fractions was achieved by differential centrifugation in order to decrease sample's complexity and

increase the proteome coverage. The analysis of membrane-enriched fractions was facilitated by membrane-associated protein solubilization in a methanol-buffered solution that enabled efficient tryptic digestion compatible with subsequent SCX-RPLC-MS [31].

The complete list of proteins identified in this analysis, as for the peptides that contributed to such identification, is provided in supplementary data (Table IV.1SD). In accordance with criteria described in *Materials and Methods*, our approach resulted in the confident identification of 535 and 1169 proteins in soluble and membrane fraction, respectively, with 222 proteins overlapping both fractions, resulting in the identification of a total of 1482 different proteins in NEC (Figure IV.1, Table IV.2SD).

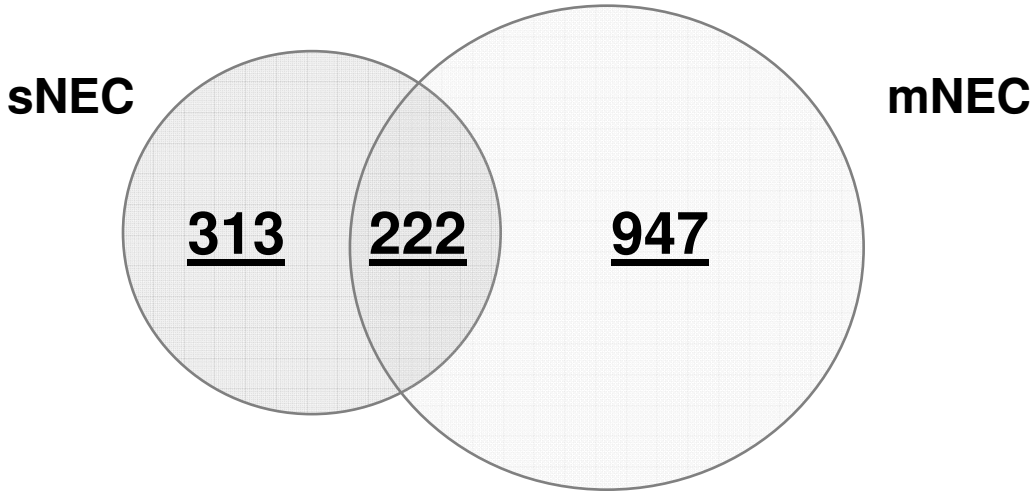


Figure IV.1: Venn diagram depicting total proteins identified in soluble (sNEC) and membrane (mNEC) fractions from nasal epithelium specimens.

By comparing our data with those previously obtained by 2D-PAGE-MALDI-TOF MS approach for NEC (136 proteins) [26,27], we observed 51 proteins commonly identified by both approaches, confirming that 2D-PAGE can still be used as a complementary tool in proteomics although with limitations.

Annotation of primary subcellular location of those 1482 proteins herein identified was performed based on information deposited in Human Proteome Reference Database (HPRD). Cytoplasmic proteins were predominant in soluble fraction (38.7%) whereas plasma membrane proteins were observed in higher percentage in membrane fraction (18.2%); their representation in overlap fraction

accounts for 5.4% of all proteins annotated in this subproteome. The term *integral to membrane* was mainly assigned to proteins allocated to the membrane fraction (6.3%; 0.5% in overlap fraction) (Figure IV.2).

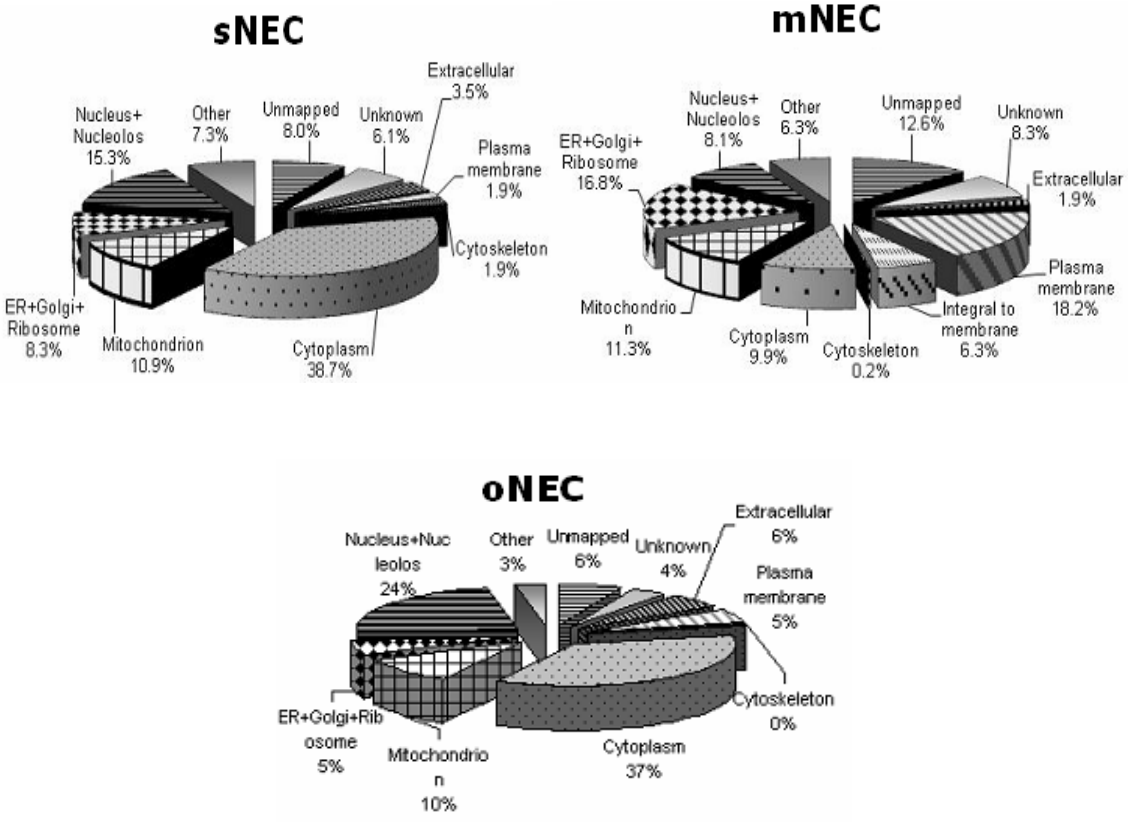
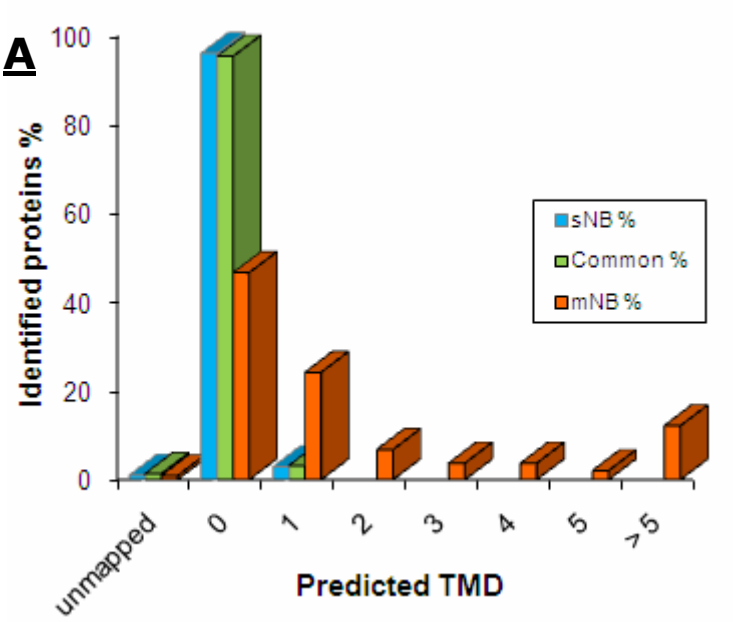


Figure IV.2: Primary localization of the identified NEC proteins from soluble **(sNEC)**, membrane **(mNEC)** and overlap **(oNEC)** fractions based on information retrieved by Human Protein Reference Database. Rare primary localization terms were grouped as *Others* representing ~ 6% of total identifications.

Putative transmembrane domains (TMD) and grand average hydropathy value index (GRAVY) were calculated, this latter as a global indicator of protein's solubility (proteins exhibiting positive GRAVY values are hydrophobic while proteins exhibiting negative GRAVY values are hydrophilic, according to [32]) (Table IV.3SD). The TMHMM analysis revealed a total of 513 proteins containing one or more putative TMDs: of those, 497 (96.8%) were found exclusively in NEC membrane fraction, corresponding to 52.5% of all proteins identified in membrane-enriched fraction (Figure IV.3A). Considering that the entire human proteome might contain about 21% of

proteins with one or more putative TMDs [34], our results represent a significant enrichment of this proteins' class. The majority of proteins with five or more predicted TMDs elicit positive GRAVY indexes (Figure IV.3B), constituting about 14% of membrane's proteome. There is, however, a significant percentage of proteins in membrane fraction with no predicted TMDs (46.7%, corresponding to 442 proteins). This observation might indicate entrapment of cytosolic proteins in membrane vesicles during lysis as well as substantial presence of membrane-associated proteins through post-transcriptional modification or protein-protein interactions. GRAVY index calculation revealed 284 out of 1482 proteins with positive GRAVY values, ranging from 0.0 to +1.2 and found exclusively in membrane fraction. In contrast, the majority of proteins identified on soluble and overlap fraction present negative GRAVY values, indicating lower hydrophobicities (Figure IV.3C). Overall, these results are in agreement with previous findings, suggesting an effective enrichment of membrane-associated proteins in membrane fraction and cytoplasmic proteins in the soluble portion.

Our datasets were also searched against human proteome database (UniProt Human) for getting attributes of NEC's identified proteins, namely existence at the protein level. In this investigation, we observed 95 proteins having only evidence at transcriptional level, three proteins are annotated as uncertain and one as *in-silico* predicted (Table IV.2.3SD).



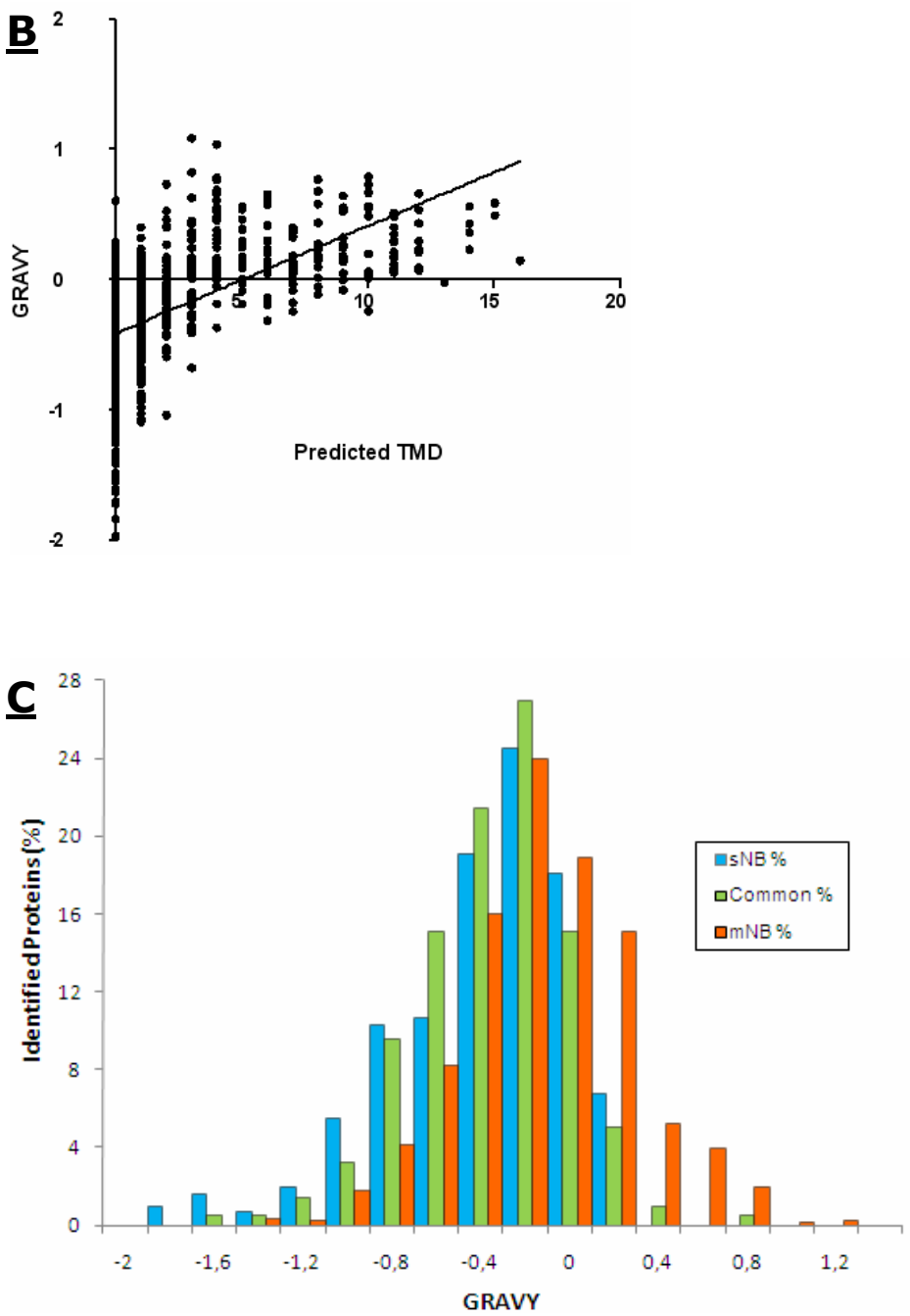


Figure IV.3: **(A)** Predicted transmembrane domains (TMD) for proteins identified in NEC calculated by TMHMM software (http://www.cbs.dtu.dk/services/TMHMM/); **(B)** Correlation between the number of predicted TMD and hydrophobic character of proteins identified in NEC; **(C)** Hydrophobicities (GRAVY values) calculated by Gene Infinity tools (http://www.geneinfinity.org/sms_proteingravy.html).

3.2 Functional Annotation of Proteins Identified by Shotgun Proteomics

Functional annotation of identified proteins was carried out using HPRD and IPA's knowledgebase. According to HPRD, an overrepresentation of biological processes such as **metabolism and energy pathways**, **cell communication and signal transduction**, **protein metabolism**, **transport**, **regulation of nucleobase** and **cellular growth and/or maintenance** was observed (Table IV.2.1SD). A considerable number of identified proteins had no annotation at all (10.7%; 158 proteins) or have missed biological function annotation (9.9%; 147 proteins). Interestingly, a higher percentage of proteins exclusively identified in membrane fraction (26%) has limited or no information available when compared with the soluble subproteome. Several proteins may be grouped in different functional families; therefore their participation in other crucial cellular processes cannot be ignored.

Among the 1458 proteins mapped by IPA's knowledgebase (Table IV.2.2SD), there is a significant overrepresentation of proteins involved in **cell death, growth and proliferation, small molecule biochemistry, cellular movement, cell-to-cell signalling and interaction, cellular assembly and organization, molecular transport, lipid metabolism and protein synthesis** ($p<0.001$) (Table IV.4SD). The subsets of 307, 938 and 213 proteins corresponding respectively to the sNEC, mNEC and oNEC subproteomes were further analysed independently by IPA to explore their relevant roles and most significant molecular bioefectors in cellular functions associated with this biological sample.

3.2.1 NEC Membrane Proteins

Among proteins exclusively identified in NEC membrane fraction, we observed an important enrichment of functions associated with **molecular transport, protein trafficking and cell-to-cell signalling/interaction** (Figure IV.4 and Table IV.4SD) bellow described in detail.

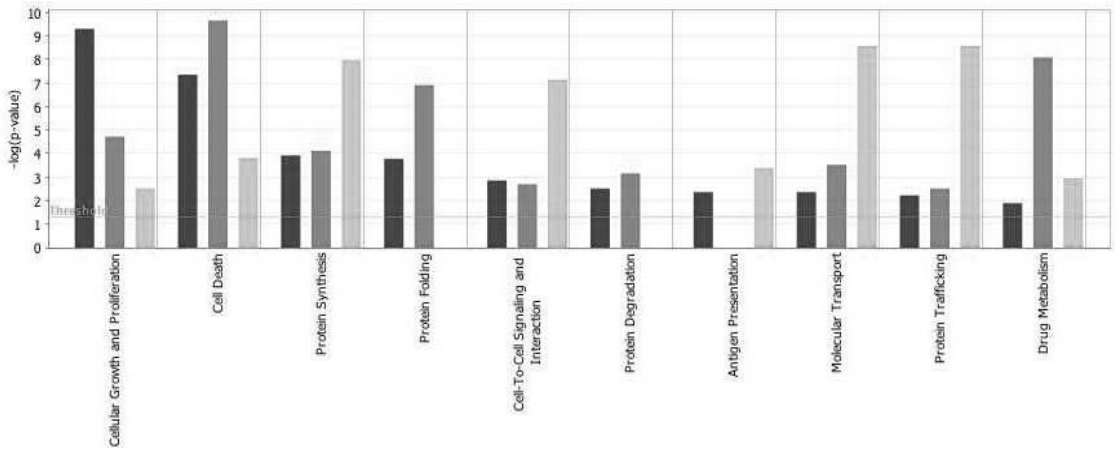


Figure IV.4: Most significant molecular and cellular functions attributed to proteins present in sNEC (dark grey), mNEC (light grey) and oNEC (black) according to IPA. The y-axis shows the - log of the *p*-value calculated based on Fisher's exact test. The dotted line represents the threshold above which there are statistically significantly more genes in a biological function than expected by chance.

3.2.1.1 Molecular transport

According to IPA analysis, the main function of proteins identified in the membrane fraction of NEC is molecular transport, mostly related to aminoacid, lipid and/or ion transportation (Table IV.4SD). Several ion channels were identified in this fraction in accordance with the major role of epithelial in ion transport processes, regulation of volume and composition of airway surface liquid (ASL), which is critical for normal lung physiology. The electrogenic Na^+/K^+-ATPase subunits (ATP1A1, ATP1B1, ATP1B3), components of a pump located at basolateral membrane that drive the absorption of Na^+ at the apical membrane through ENaC [35], were identified. Na^+ absorption generates a transepithelial electrical potential difference that leads to transepithelial Cl^- absorption to maintain the electroneutrality of the ion transport process. Cl^- first enters the cell through a basolateral Na^+-K^+-Cl^- co-transporter (NKCC2), which was identified in our data. Cl^- is then secreted at the apical membrane via, for example, the cystic fibrosis transmembrane conductance regulator (CFTR). The electrical driving force for Cl^- secretion is also supplied by basolateral K^+ channels [36]. Curiously, despite we can assess CFTR in human NEC by immuncoytochemistry [29,37,38] neither CFTR, ENaC or basolateral K^+ channels were uncovered from the MS data based on the two unique peptides in three independent runs' restriction criteria used in our analysis. However, lowering the restriction criteria from two to one peptide, CFTR was identified in four independent experiments by a total of five unique peptides,

all corresponding to cytoplasmic domains of the CFTR. Technical difficulties associated with high hidrophobicity and glycosilation of CFTR together with low expression levels of CFTR in NEC seem to be strong arguments to explain our observation [39,40]. Similarly, basolateral K^+ channels such as KCNN1, KCNN4, KCNJ5, KCND1 and KCNAB1 were also identified according to this less conservative analysis, probably reflecting the reduced abundance of these proteins in the cells, regardless their imperative role for cellular function.

A subunit of the electrogenic H^+/K^+-ATPase (ATP12A) was also identified here, which, along with Na^+/K^+-ATPase, is responsible for maintenance of osmotic balance and intracellular ionic composition. Absorption or secretion of water is driven by osmosis, determined by transepithelial NaCl movement, and occurs through the lipid layer and through aquaporins [41,42]. Here, AQP5 is the only molecule identified representing water movement across nasal epithelium.

Our data also revealed the presence of the Na^+/H^+ exchanger NHE1 and the bicarbonate transporter NBC1, which are involved in the maintenance of epithelial acid-base balance. The countertransport of H^+ in an electroneutral manner ($1Na^+$:$1H^+$ stoichiometry) by NHE exchangers, propelled by the inward Na^+ gradient established by plasma membrane Na^+/K^+-ATPase pumps, allows extrusion of H^+ excess (acid equivalents) accumulated by cellular metabolism. Additionally, bicarbonate cotransporters allow extrusion of $HCO3^-$ ($1Na^+$:$1HCO3^-$ stoichiometry), a cellular biological buffer that plays an important role on solubilization of macromolecules such as mucins and digestive enzymes in secreted fluids. Thus, the activity of those proteins are crucial not only to fine control of intracellular pH and but also to physiological processes such as cell volume control, systemic electrolyte, acid-base fluid volume homeostasis and epithelia protection from injury [43,44].

Other ATPases were represented in this fraction, namely Ca^{2+} pumps. Three Ca^{2+}-ATPases have been described in the cells of higher animals. They are located in the membranes of endo(sarco)plasmic reticulum, including the nuclear envelope (SERCA pump), the Golgi network (the SPCA pump) and plasma membrane (PMCA pump) [45]. Each pump is a product of a multigene family, the number of isoforms being further increased by alternative splicing of the primary transcripts. Here, several components of those three $Ca2^+$ pumps were identified: ATP2A1, ATP2A2 and ATP2A3 from SERCA pump; ATP2C1 from SPCA pump; and ATP2B1, ATP2B4 from

PMCA pump [45]. Components of V-type H^+-ATPase were also revealed, namely two subunits belonging to the membrane-embedded V0 complex (ATP6V0A4, ATP6V0C) and two subunits belonging to the cytoplasmic V1 complex (ATP6AP1, ATP6V1E1), responsible for H^+ translocation across the membrane and mediation of ATP hydrolysis. Firstly associated with endosomal membranes, this pump was also found in plasma membranes where the proton pump energizes transport across cell membranes and entire epithelia [46]. The Cu^{2+}-ATPase 7B (ATP7B), also observed in our data, is responsible for the delicate balance of Cu^{2+} in the cells and throughout the whole body in association with ATP7A. Its absence or malfunction leads to severe human genetic disorders, Menkes and Wilson diseases. ATP7B is primarly expressed in the liver but it was also observed at lower levels in brain, heart and lungs [47]. Combined data from various reports on polarized hepatic cells suggest a continuous recycling of ATP7B-containing vesicles between sub-apical compartment and the apical membrane for Cu^{2+} release as its levels increased in the cell [47].

Among the identified proteins, it should be highlighted the identification of membrane proteins responsible for drug transport belonging to the ATP-binding cassette (ABC) transporter C family (ABCC1, ABCC3, ABCC4). In particular, ABCC1 (also known by MRP1, Multidrug Resistance Protein 1) is highly expressed mainly at the basolateral side of bronchial epithelial cells, namely ciliated and goblet cells. This is consistent with our data where a high number of unique peptides were detected for this protein (Table IV.1SD). MRP1 confers resistance to several chemotherapeutic agents including vincristine, daunorubicin and methotrexate. Physiological substrates for MRP1 are *e.g.*, leukotriene C4 (LTC4) and glutathione disulfide. Interestingly, these substrates play an important role in lung physiology with respect to inflammation and oxidative stress [48]. ABCC3 (or MRP3) gene is the closest MRP1 homologue. Its product is involved in resistance against the anti-cancer drugs etoposide, teniposide and, at higher concentrations, also to methotrexate. The physiological function for ABCC4 (or MRP4) is still largely unknown. They may serve as an efflux pump of nucleosides cAMP and cGMP at low affinity, likely in a GSH-independent manner, having thus different substrate's specificities. In summary, expression levels of these transporters at airway epithelium, in particular NEC should be crucial for detoxification and might have implications in therapeutic strategies for respiratory diseases.

3.2.1.2 Protein Trafficking

IPA also ascribed a large number of proteins from membrane fraction to vesicular transport. From a total of 56 proteins, several are known to be involved in SNARE protein-mediated membrane trafficking, responsible for docking and fusion of vesicles to targeted membrane. Associated-molecular machinery identified in this fraction includes vesicle-trafficking protein SEC22B and SEC23-interacting proteins, STX7, STX8 and STX18, VAMP2, VAPA, USE1, SNAP23 and GOSR1 proteins [49]. In this functional group, it is also observed the secretory carrier membrane proteins 1, 2 and 3 (SCAMP1, SCAMP2 and SCAMP3), α- and γ1-adaptin (AP1G1 and AP2A1, respctively), all components of adaptor protein complexes involved in cargo selection and vesicle formation associated to clathrin-dependent endocytosis [50-52]. Proteins belonging to GTPase family, representing regulators of eukaryotic vesicular membrane traffic [53] such as RAB10, RAB13, RAB6A and RAB7A, were also grouped here.

Concordantly, IPA revealed caveolar-mediated endocytosis as a canonical pathway enriched in membrane fraction. Caveolae are plasma membrane invaginations characterized by flask-shape morphology, enriched in cholesterol and glycosphingolipids. They are proposed to be important in lipid transport, acting as scaffolding or organizing platforms for signalling events that include endothelial nitric oxide synthase (eNOS) activation and recruitment of protein kinase C isoforms to the plasma membrane [54]. Two proteins, flotilin 1 and 2 (FLOT1 and FLOT2) known as structural molecules for caveolar support were also identified. Curiously, Caveolin-1 (CAV1) was only observed after opening previously described criteria. According to some evidences, flotilins are involved in an alternative noncaveolin-dependent mechanism for structural support of caveolae [55,56]. Whether our observation reflects the predominance of flotilins over CAV1 in NEC or if it is an artefact of sample preparation remains to be elucidated.

Evidences show a link between caveolar-mediated endocytosis and regulation of cell adhesion [57], through internalization of integrins (ITGA2, ITGA3, ITGA6, ITGAM, ITGAV, ITGB1, ITGB2 and ITGB4), also grouped here. The same was observed for several members of small GTPase superfamily, RAB5A, RAB5B and RAB5C, which seem to have an important role in caveolae dynamics [58].

3.2.1.3 Cell-to-cell signalling and interaction

Significant number of proteins associated with cell-to-cell signalling/interaction, were identified and include intercellular adhesion molecules which make part of tight (CLDN1, CLDN3, F11R), adherens (CDH1, CTNNA1, CTNNB1, CTNND1), desmosome (PKP2, DSG2, DSG3) and hemidesmossome junctions (ITGA2, ITGB1, ITGAV, ITGA6, ITGA3, ITGB2, ITGB4) [59]. These structures are located at plasma membrane and have an important function of establishing physical interactions between adjacent cells, maintaining integrity and impermeability of airway epithelia [60,61]. In addition, a substantial number of proteins are involved in receptor activity which is compatible with the important role of epithelial cells in sensing changes in its environment [3]. This group includes proteins belonging to the small GTPase superfamily (RAB13, RAB21, RAP1A, RAC1), tyrosine protein kinase family (EGFR, IGF1R, ERBB2, EPHA2) and tyrosine protein phosphatase family (PTPRC, PTPRJ). Some molecules involved in immune response were also grouped here. It is the case of HLA-A, HLA-DRA proteins, constituents of the MHC class I and II, respectively, the antigen binders IGHG1, IGKC, IGHM, and leukocyte surface markers such CD47, CD9, CD59a and CD14. The latter marker was already described as cooperating with MD-2 and TLR4 to mediate the innate immune response to bacterial LPS [3,62]. Other LPS-binding proteins such as neutrophil defensin, calgranulin A and B, cathelicidin antimicrobial peptide and bactericidal permeability-increasing protein were also identified, mostly in NEC soluble fraction, described as follows.

Molecules involved in immune response can arise from several sources, including inflammatory cells resident in the epithelium. In accordance with current knowledge, their interaction with respiratory epithelia modulates innate and acquired immune response against microbial infections and inhaled toxic pollutants [16,18].

3.2.2 NEC Cytoplasmic and/or Soluble Proteins

Biological functions enriched in soluble fraction subproteome account for **cell death, drug metabolism and protein folding**, typical of cytoplasm/intracellular compartments (Table IV.4SD). The most relevant functions are discussed below.

3.2.2.1 Cell death

Cell death was found by IPA analysis as the main molecular and cellular function of proteins identified in the soluble fraction of NE (Figure IV.4). Exploring this group of proteins we observed different pathways contributing for this function such as **oxidative stress response mediated by Nrf2, Myc mediated apoptosis signalling and also metabolism namely glycolysis/gluconeogenesis and citrate cycle**.

3.2.2.2 Oxidative stress response

Glutathione (GSH) metabolism is one of the enriched functions retrieved by IPA in the soluble fraction for which several identified proteins make a noteworthy contribution (GSR, GSTP1, GSTA1, GCLC, IDH3A, GSTO1, PRDX6, IDH1) as well as NRF2-mediated oxidative stress response [63,64] (AKR7A2, GSTP1, SOD1, GSTA1, NQO1, GCLC, GSTO1, GSR, AKR1A1, SOD2, ERP29, STIP1, CAT, TXN) and free radical scavenging (CLIC6, DNL, HNRNPA3, ILF2, LGALS3BP, PDIA6, PPIA, PRB3, PRDX1, PRDX6, S100A8, TALDO1, TST, UBXN, ALB, ALDH2, GPX1, HSD17B10, HSPB11), especially reduction and catabolism of hydrogen peroxide (MPO, GPX1, PRDX1, PRDX2, PRDX3, PRDX5). GSH is one of the most powerful endogenous antioxidants, protecting molecules by conjugation with hydrogen peroxide to yield water and oxygen via glutathione peroxidase (GPX1). Oxidized glutathione (GSSG) is then regenerated to GSH by NADPH via glutathione reductase (GSR). NFR-2 [Nuclear factor (erythroid-derived 2)-like 2] is a transcription factor master regulator of the antioxidant response [63-65] by inducing expression of genes involved in combating oxidative stress and activating body's own protective response. Examples of these are glutathione-S-transferase (GST) family, a group of cytosolic, mitochondrial and microsomal enzymes that catalyze the conjugation of GSH with both endogenous and xenobiotic electrophiles, whose members GSTA1, GSTK1, GSTO1 and GSTP1 were identified here, and glutamate-cysteine ligase (GCLC), enzyme

involved in the rate-limiting step in the synthesis of GSH, was also found. GST and GCLC are induced by NFR-2 activation and represent an important route to eliminate potentially harmful and toxic compounds while balancing redox state of a cell [66,67]. This is in agreement with higher proportion of identified proteins involved in drug and xenobiotics metabolism (ALDH5A1, ALDH7A1, ALDH9A1, CAT, CES1, CES2, GCLC, GSTA1, GSTO1, GSTP1, HSP90AB1, NQO1, PPP2R1A). All these proteins are involved in the maintenance of an adequate oxidant/antioxidant balance in NEC, playing a decisive role in ameliorating the oxidative stress and protect the airway against hostility and injury promoted by the environment. Accumulation of ROS with concomitant attack to cellular membranes might compromise cellular structure maintenance and tissue integrity, therefore combating oxidative stress is a critical and vital function in this tissue.

A slight enrichment of DNA replication, recombination and repair functions in soluble fraction was also observed (PTMS, GSTP1, UBE2N, GCLC, CBX3, TYMP, SET, PRDX6, ACIN1, NCL, SOD2, CAST, APEX1, RAD23B, NAP1L4, CALR, SOD1, CFL1, YWHAE, PPIA, ERH, NQO1, VIM, HMGN1, DDB1, ASAH1, NASP, PRTN3, XRCC6, CAT, CYCS, HMGN2, TMPO, NPM1), which, in conjugation with oxidative stress response machinery, may orchestrate a particular proteinaceous environment propitious to cellular division in accordance with the rapid turnover and regeneration of NE.

3.2.2.3 Glucose metabolism and reducing equivalents production

A high proportion of proteins involved in glycolysis/gluconeogenesis (ADH1C, PGM1, TPI1, ALDH9A1, LDHB, AKR1A1, GPI, ADH7, ENO1, PGAM1, DLD, FBP1, ALDH5A1, ALDOC, ALDH7A1), pentose phosphate pathway (GPI, HIBADH, PGLS, TALDO1, PGM1, FBP1, ALDOC) and citrate cycle (SUCLA2, CS, SUCLG2, SUCLG1, DLD, IDH3A, MDH1, FH, ACO1, IDH1), processes taking place in the cytosol and mitochondrial matrix, were predominantly allocated to the soluble fraction. (Table IV.4SD). The latter mentioned processes are closely related to the unquestionable importance and function of NE as the first barrier between environment and organism's interior. This interface is constantly being exposed to physical and chemical aggressions resulting in increased wound healing and remodelling [16,18,68,69]. It is therefore expected that this epithelium has increased demands of

high energy compounds (ATP and NADH), reducing equivalents (NADPH) and pentoses, molecules derived from the mentioned processes, while maintaining a constant reservoir of glucose to guarantee proper physiological functions and assure tissue integrity and structure. Of particular interest is the generation of pentoses for the synthesis of nucleotides and nucleic acids for consequent use in genomic material processing and NADPH which main functions account for reductive biosynthesis (*e.g.* fatty acids synthesis) and protect cells from oxidative stress by glutathione reduction [70].

Some abundant proteins associated with metabolism and energy pathways (PKM2, ALOX15, HK1, HADHB, LDHA, HADHA, PGK1, FDXR, ATP5B, ATP5A1), regulation of gene expression and nucleobase, nucleoside, nucleotide and nucleic acid metabolism (ILF3, PML, DDX3X, EWSR1, LGALS3, XRCC5) although predominately found in the soluble fraction was also observed in membrane fraction.

3.2.2.4 Protein folding and degradation

Protein folding and degradation also appear as processes overrepresented in soluble fraction. These processes occur naturally in the cytosol associated with ER, Golgi and transport vesicles constituting pathways of proteins' processing and maturation until a final functional form.

Proteins such as CALR, UGGT1, ERP29, ERP44, HSP90AB1, HSP90AA1, PDIA6, PPIA, HSPE1, ERO1L, TXN, ST13, HSPA5, HSPA8, HSPD1 and RUVBL2 are mainly assigned to folding while ALDH1A1, ALDH3A1, MPO, HSP90B1 and HSPA5 strongly contribute to proteins' turnover mechanisms. Several of the identified proteins present chaperone activity ensuring proper proteins' homeostasis and upholding of a functional tissue.

Post-translational modifications are the late stage phenomena occurring during protein synthesis and influence protein's conformation and therefore its function. Several identified proteins (PEBP1, PRDX4, UBE2N, GNB2L1, ST13, SET, PRDX6, HSPA4, SOD2, HSP90AB1, PARK7, HSPE1, SERPINA1, TXN, CAST, PPP1CA, ALDH5A1, ALDOC, CALR, P4HB, PPIA, OAT, ERO1L, DDB1, PPP2R1A, UGGT1, ERP29, ERP44, PDIA6, CAT, PSAP, GLUL, CAPN2, UBA1) are closely related to this event. The modifications highlighted by the stated proteins are mainly aminoacid-based,

like auto-oxidation of cysteines and tyrosines by CP, ALB and MPO, respectively, deamination of glutamine acid by GLUD1 or exposure of lysines by ANXA2 while proteins' homotetramerization is the result of the action of ALDH1A1, DECR1 and HSD17B10. However, post-translational modifications of aminoacids extend the range of protein's functions by attaching to it other biochemical functional groups, changing its chemical nature or by making structural rearrangements.

3.2.2.5 Cytoskeleton

A predominance of actins and actin-related proteins were found in the soluble fraction. This is illustrated by the identification of ACTN1, ACTR2, ACTR3, ARPC2, ARPC3 and ARPC4 but also CFL1, GSN, MSN, MYL6, MYL12A, PPP1CA, TMSB4X, VCL and ARHGDIA.

Most of these cytoskeleton constituents are involved in cellular assembly, organization and Rho-dependent movement, biological processes also enriched in the soluble fraction subproteome. In fact, members of the Rho GTPase family have been shown to regulate many aspects of intracellular actin dynamics [71], acting like molecular switches of cell proliferation, apoptosis, cell polarity or vesicular trafficking. They also participate in the formation of the lamellipodia and filopodia by encouraging actin retraction [72], fundamental to cellular movement and wound healing. Corroborating this outcome was the identification of several subunits of the ARP2/3 complex (ARPC2, ARPC3, ARPC4, ACTR2, ACTR3), proteins present at the microfilaments junctions that help to create the actin meshwork [73]. Observation of the mentioned proteins substantiates their part in the maintenance of an architectural structure that fits the function of the epithelium.

Some other components of cytoskeleton (PFN1, COL1A2, ACTB, ACTN4, DSP, KRT17, KRT10, LMNA, YH11, PLEC1), and cell communication processes (ANXA11, ANXA2, ANXA1, S100A11, LCP1, JUP, RAB1A) found in the soluble fraction were also identified in the membrane fraction, reflecting their close association with membranes or membrane proteins.

3.3 Protein Abundances according to emPAI and % mol

The characterization of proteins' abundance in a particular cell/tissue may provide important information about its functional contribution to the organism physiology. Qualitatively, some parameters such as hit rank, score and number of peptides *per* protein, *i.e.*, integrated peptide ion count measurements [74], can be considered as indicators of protein abundance in a sample [75]. Recently developed, identification-based algorithms that include emPAI have shown a high correlation with the actual protein amount in complex mixtures with a wide dynamic range [33]. Although emPAI's accuracy is inferior to absolute quantification using synthesized peptide standards, it is effective in providing the information of protein abundances within the proteome of interest. We used emPAI to estimate protein content in NEC, expressed in molar percentages for each identified protein (% mol), highlighting the top 100 most abundant and scarce proteins in each subproteome (Table IV.5SD).

3.3.1 Top 100 Most Abundant Proteins

The top 100 most abundant proteins constitute 49.4% and 39.8% of the total proteins identified in membrane and soluble fraction, respectively (Table IV.5.3SD). Among the top 100 proteins, mitochondrial and ribosomal proteins are highly represented. Several subunits of cytochrome c oxidase, the terminal oxidase in mitochondrial electron transport (MT-CO2, COX4I1, COX5B, COX6C), and subunits of the mitochondrial membrane ATP synthase (ATP5F1, ATP5H, ATP5I, ATP5J2, ATP5L), which are involved in energy production and ATP metabolism, were detected in high abundance. Structural constituents of ribosome, namely several members of the 40S and 60S subunits of the ribosome (RPS14, RPS16, RPS17, RPS18, RPS24, RPS25, RPS3, RPS9, RPL18, RPL19 and RPL7) were also stressed up.

Keratins (KRT10, KRT18, KRT19, type I keratins and KRT1, KRT5, KRT7, KRT8, type II keratins) and tubulins (TUBB2A and TUBA1A), among others (PFN1, ACTA2, GSN, TMSB4X) were also found in a high abundance both in membrane and soluble fractions. High abundance of keratins is in agreement with its known importance as structural stabilizers of epithelial cells given its polimerization into keratin filaments. Intracellularly, they braid the nucleus, span through the cytoplasm and are attached to the cytoplasmic plaques of the desmosomes. Thus, it seems that they are

inherent part of the continuum of stability from the single cell to the tissue formation [76]. The identified tubulin isoforms, TUBB2A and TUBA1A, were described as having additional role in airway epithelium besides being important components of microtubules and building blocks of cilia of tall columnar cells [77].

Identification of LPLUNC1 as one of the most abundant proteins in NEC membrane fraction seems relevant for innate immunity in upper airway. This protein belongs to the PLUNC family which are structural homologues of LPS-binding protein and the bacterial permeability–increasing protein, both of which are known mediators of host defence against gram-negative bacteria [78]. Additionally, other antimicrobial components such as cathelicidin antimicrobial peptide (CAMP) and cathepsin G were also included here.

Proteins associated with protein folding and processing such as ANXA2, HSPB1, PDIA3, HSP90AA1, P4HB, HSPA9, PPIA, HSP90B1, HSPA8, HSPA1A, HSPD1 and PDIA4 are also among the top 100.

Another key function that is associated with the group of abundant proteins is metabolism and detoxification of xenobiotics by cytochrome P450 (ALDH2, GSTP1, ADH7, GSTA1, ALDH3A1, GSTK1). This occurs in tight conjugation with antioxidant response of the cells and glutathione association (GSTP1, GSTA1, IDH2, GSTK1) to eliminate potentially harmful compounds that might contribute to destabilization of the tissue.

Several top 100 abundant proteins have been described as associated with development, progression and or metastasis of lung cancer (AGR2, ALDH3A1, ALDOA, GSN, HSP90AA1, HSP90B1, KRT5, KRT19, PPIA, S100P, SERPINB3, TMSB4X, TPI1, TUBA1A, VIM, YWHAE, S100A4, CTSB) and other respiratory disorders (ACTA2, GAPDH, HBB, HSP90B1, S100A4, S100A9, S100P, TALDO1, TXN). Two key proteins previously described in a cystic fibrosis mouse model by our group [79], retinal dehydrogenase 1 (ALDH1A1) and aldehyde dehydrogenase (ALDH2), along with ADH7, are crucial enzymes in retinoic acid metabolism, a hormone known as important regulator of organ development and homeostasis, including the lung [80]. Expression of ALDH1A1 in airway epithelium has already been confirmed by *in situ* hybridisation of embryonic lung [81].

3.4 NE proteome relevance to respiratory diseases

The comparison of NEC proteome herein obtained (1482 proteins) with the one previously reported for mainstream bronchus epithelium (859 proteins) also harvested by brushing [82] revealed that there are about 517 proteins commonly expressed by both airways epithelia (Figure IV.5 and Table IV.6SD). Despite those authors have used 1D-PAGE-LC-MS/MS, a technique adequate for membrane proteins' analysis, the nature of these overlapping proteins is mainly soluble probably due to the fact that we have used prefraccionation strategies for higher enrichment of membrane proteins. The major molecular function associated to the overlapping portion is catalytic activity, including oxireductase and hydrolase activities.

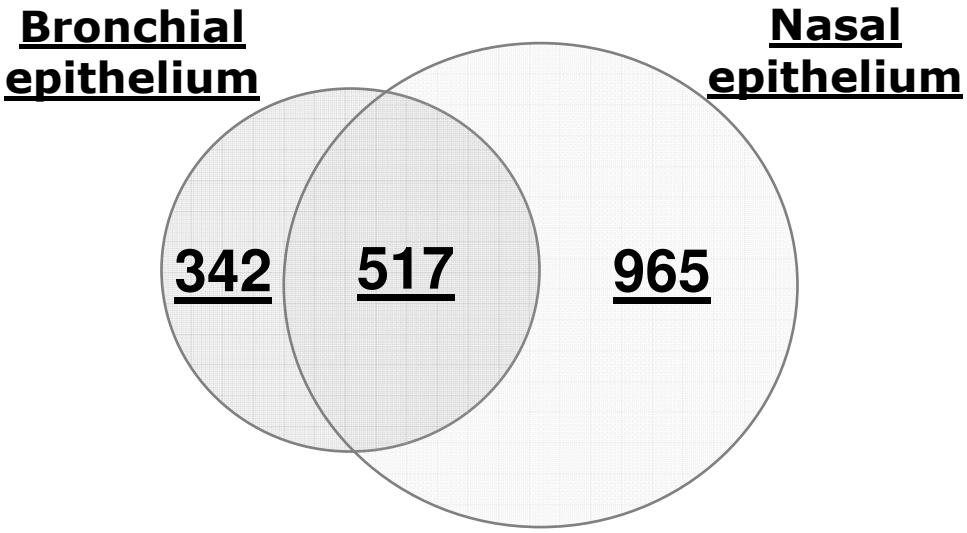


Figure IV.5: Venn diagram depicting a comparison of proteins identified in nasal epithelium in our study with those identified in bronchial epithelium by Steiling *et al.* 2009. Comparative analysis was performed by Ingenuity Pathways Analysis software.

Among these, 142 proteins are annotated by HPRD as being expressed in respiratory epithelium that includes not only nasal and bronchial epithelium but also lung epithelium and bronchoalveolar fluid (BALF). For example, PLUNC, which we described above as one of the top 100 most abundant proteins, polymeric immunoglobulin receptor and aldehyde dehydrogenase 3B1 were described by Steiling *et al.* [82] and observed within bronchial epithelial cells, airway epithelia and lung, respectively, accordingly to HPRD. Proteins such as CD44, HLA-DRA, MUC1 and

AQP5, described in the present work (mNEC proteome) but not in Steiling's work, are also annotated in HPRD as expressed in bronchial epithelial cells.

Altogether, these results corroborate previous idea that NEC can reflect airway molecular environment and thus be used as a complementary sample in biomarker discovery for pulmonary diseases. In fact, about 141 proteins herein identified have been reported as associated with the development and/or progression of several respiratory diseases (Table IV.7SD). Lung cancer appears as the pathology with more associated proteins (75 out of the 141). Other diseases emerged from IPA's analysis namely severe acute respiratory syndrome (SARS) (ACSL1, ACTN1, ANPEP, BPI, CAMP, CD9, CD63, CLU, DEFA1, FLOT2, GAPDH, GLUL, GSTO1, HBG1, HIST1H1C, HP, ITGAM, LCN2, LTA4H, LTF, MNDA, MTX1, PDXK, RAB13, S100A9, S100P, SERPINA1, TALDO1, TCIRG1, TKT, TUBB2C) or pneumonitis (ACO2, ACTB, ALB, ALDH2, ALDOA, ANXA3, CAMP, ENO1, HSPA5, HTT, MYH9, NOS2, P4HB, PDIA3, PRDX1, PRDX6, SELENBP1, STAT3, TKT, TOP2B, TPI1, TUBB). Two proteins, NOS2 and TPI1, have been associated with asthma while MPO and LTF, as mentioned before, have been involved in the development of Cystic fibrosis [83,84]. In addition, recent results from our group and others also showed that ENO1 and SELENBP1 are in somehow implicated in the progression of Cystic Fibrosis.

4. SUMMARY

The present study was carried hypothesizing that nasal epithelia reflects lower airways' events and thus be a useful alternative specimen to highly invasive biopsies for biomarker discovery-based studies on lung diseases. Although nasal fluid, bronchoalveolar lavage and sputum have been extensively used to assess airway physiology, they mainly provide secreted and tissue leakage molecules that may not effectively reflect cellular events. Here, we delivered to most complete to date proteome profiling of NEC by using an improved subcellular fractionation method followed by a high sensitive multidimensional protein identification technology and unravelled mechanisms and processes intimately correlated with normal and pathological lung function. Most of the identified NEC proteome overlaps with that previously described for bronchial epithelia providing extra consistency on our contend about the usefulness of these cells in lung research.

The results achieved here can serve as basis for better understanding of molecular features underlying chronic respiratory diseases using non-invasive and easily collectable specimens that are physiologically and molecularly similar to those on lower airways.

REFERENCES

[1] Crystal RG, Randell SH, Engelhardt JF, Voynow J, Sunday ME. Airway epithelial cells: current concepts and challenges. Proc Am Thorac Soc 2008; 5:772-7.

[2] Pier GB. Airway epithelial (nasal) cell monolayers used to study Pseudomonas aeruginosa invasion are hyperpolarized and not representative of the human airway epithelium. Infect Immun 2006; 74:7043; author reply 7043-4.

[3] Vroling AB, Fokkens WJ, van Drunen CM. How epithelial cells detect danger: aiding the immune response. Allergy 2008; 63:1110-23.

[4] Gaga M, Vignola AM, Chanez P. Upper and lower airways: similarities and differences. Eur Respir Mon 2001; 18:1-15.

[5] Mygind N, Pedersen M, Nielsen M. Morphology of the upper airway epithelium. in: D.F. Proctor, and I. Andersen, (Eds.), The Nose, Upper airway physiology and the atmospheric environment, Elsevier Biomedical Press, Amsterdam, 1982, pp. 15-45.

[6] Ayers MM, Jeffery PK. Proliferation and differentiation in mammalian airway epithelium. Eur Respir J 1988; 1:58-80.

[7] Bolduc P, Jones R, Reid L. Mitotic activity of airway epithelium after short exposure to tobacco smoke and the effect of the anti-inflammatory agent phenylmethyloxadiazole. Br J Exp Pathol 1981; 62:461-8.

[8] Jeffery PK. Microscopic structure of normal lung. in: R.A.L. Brewis, B. Corrin, D.M. Geddes, and G.J. Gibson, (Eds.), Respiratory Medicine, Saunders, London, 1995, pp. 25-39.

[9] Boucher RC, Stutts MJ, Knowles MR, Cantley L, Gatzy JT. Na+ transport in cystic fibrosis respiratory epithelia. Abnormal basal rate and response to adenylate cyclase activation. J Clin Invest 1986; 78:1245-52.

[10] Slavin RG. The upper and lower airways: the epidemiological and pathophysiological connection. Allergy Asthma Proc 2008; 29:553-6.

[11] McDougall CM, Blaylock MG, Douglas JG, Brooker RJ, Helms PJ*, et al.* Nasal epithelial cells as surrogates for bronchial epithelial cells in airway inflammation studies. Am J Respir Cell Mol Biol 2008; 39:560-8.

[12] Bonestroo HJ, de Winter-de Groot KM, van der Ent CK, Arets HG. Upper and lower airway cultures in children with cystic fibrosis: Do not neglect the upper airways. J Cyst Fibros 2010; 9:130-134.

[13] Hurst JR, Wilkinson TM, Perera WR, Donaldson GC, Wedzicha JA. Relationships among bacteria, upper airway, lower airway, and systemic inflammation in COPD. Chest 2005; 127:1219-26.

[14] Zulianello L, Canard C, Kohler T, Caille D, Lacroix JS*, et al.* Rhamnolipids are virulence factors that promote early infiltration of primary human airway epithelia by Pseudomonas aeruginosa. Infect Immun 2006; 74:3134-47.

[15] Togias A. Systemic cross-talk between the lung and the nose. Am J Respir Crit Care Med 2001; 164:726-7.

[16] Thompson AB, Robbins RA, Romberger DJ, Sisson JH, Spurzem JR, *et al.* Immunological functions of the pulmonary epithelium. Eur Respir J 1995; 8:127-49.
[17] Knight DA, Holgate ST. The airway epithelium: structural and functional properties in health and disease. Respirology 2003; 8:432-46.
[18] Swindle EJ, Collins JE, Davies DE. Breakdown in epithelial barrier function in patients with asthma: identification of novel therapeutic approaches. J Allergy Clin Immunol 2009; 124:23-34; quiz 35-6.
[19] Shi W, Bellusci S, Warburton D. Lung development and adult lung diseases. Chest 2007; 132:651-6.
[20] Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, *et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989; 245:1066-73.
[21] Pacheco P, Detection rate of CF causing CFTR mutations in Portugal. in: WHO, (Ed.), The Molecular Genetic Epidemiology of Cystic Fibrosis, Human Genetics Programme, Chronic Diseases and Health Promotion, 2004.
[22] Cystic Fibrosis Foundation, Cystic Fibrosis Mutation Database, Bethesda, 1989.
[23] Ollero M, Brouillard F, Edelman A. Cystic fibrosis enters the proteomics scene: new answers to old questions. Proteomics 2006; 6:4084-99.
[24] Ratjen F. Update in Cystic Fibrosis 2008. Am J Crit Care Med 2009; 179:445-448.
[25] Lemanske RF, Jr., Busse WW. Asthma: clinical expression and molecular mechanisms. J Allergy Clin Immunol 2010; 125:S95-102.
[26] Roxo-Rosa M, da Costa G, Luider TM, Scholte BJ, Coelho AV, *et al.* Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006; 6:2314-25.
[27] Lee JY, Byun JY, Lee SH. Proteomic analysis of normal human nasal mucosa: establishment of a two-dimensional electrophoresis reference map. Clin Biochem 2009; 42:692-700.
[28] Passalacqua G, Ciprandi G, Canonica GW. The nose-lung interaction in allergic rhinitis and asthma: united airways disease. Curr Opin Allergy Clin Immunol 2001; 1:7-13.
[29] Penque D, Mendes F, Beck S, Farinha C, Pacheco P, *et al.* Cystic fibrosis F508del patients have apically localized CFTR in a reduced number of airway cells. Lab Invest 2000; 80:857-68.
[30] Dacie JV, Lewis S. Practical Haematology. Edinburgh:1984.
[31] Blonder J, Chan KC, Issaq HJ, Veenstra TD. Identification of membrane proteins from mammalian cell/tissue using methanol-facilitated solubilization and tryptic digestion coupled with 2D-LC-MS/MS. Nat Protoc 2006; 1:2784-90.
[32] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982; 157:105-32.
[33] Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, *et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 2005; 4:1265-72.
[34] Xiao Z, Blonder J, Zhou M, Veenstra TD. Proteomic analysis of extracellular matrix and vesicles. J Proteomics 2009; 72:34-45.
[35] Jasti J, Furukawa H, Gonzales EB, Gouaux E. Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH. Nature 2007; 449:316-23.

[36] Bardou O, Trinh NT, Brochiero E. Molecular diversity and function of K+ channels in airway and alveolar epithelial cells. Am J Physiol Lung Cell Mol Physiol 2009; 296:L145-55.
[37] Carvalho-Oliveira I, Efthymiadou A, Malho R, Nogueira P, Tzetis M*, et al.* CFTR localization in native airway cells and cell lines expressing wild-type or F508del-CFTR by a panel of different antibodies. J Histochem Cytochem 2004; 52:193-203.
[38] Davidson H, McLachlan G, Wilson A, Boyd AC, Doherty A*, et al.* Human-specific cystic fibrosis transmembrane conductance regulator antibodies detect in vivo gene transfer to ovine airways. Am J Respir Cell Mol Biol 2006; 35:72-83.
[39] Chang XB, Hou YX, Jensen TJ, Riordan JR. Mapping of cystic fibrosis transmembrane conductance regulator membrane topology by glycosylation site insertion. J Biol Chem 1994; 269:18572-5.
[40] Wright JM, Merlo CA, Reynolds JB, Zeitlin PL, Garcia JG*, et al.* Respiratory epithelial gene expression in patients with mild and severe cystic fibrosis lung disease. Am J Respir Cell Mol Biol 2006; 35:327-36.
[41] Verkman AS. Physiological importance of aquaporins: lessons from knockout mice. Curr Opin Nephrol Hypertens 2000; 9:517-22.
[42] Tradtrantip L, Tajima M, Li L, Verkman AS. Aquaporin water channels in transepithelial fluid transport. J Med Invest 2009; 56 Suppl:179-84.
[43] Orlowski J, Grinstein S. Diversity of the mammalian sodium/proton exchanger SLC9 gene family. Pflugers Arch 2004; 447:549-65.
[44] Dorwart MR, Shcheynikov N, Yang D, Muallem S. The solute carrier 26 family of proteins in epithelial ion transport. Physiology (Bethesda) 2008; 23:104-14.
[45] Brini M, Carafoli E. Calcium pumps in health and disease. Physiol Rev 2009; 89:1341-78.
[46] Beyenbach KW, Wieczorek H. The V-type H+ ATPase: molecular structure and function, physiological roles and regulation. J Exp Biol 2006; 209:577-89.
[47] La Fontaine S, Mercer JF. Trafficking of the copper-ATPases, ATP7A and ATP7B: role in copper homeostasis. Arch Biochem Biophys 2007; 463:149-67.
[48] van der Deen M, de Vries EG, Timens W, Scheper RJ, Timmer-Bosscha H*, et al.* ATP-binding cassette (ABC) transporters in normal and pathological lung. Respir Res 2005; 6:59.
[49] Yoshizawa AC, Kawashima S, Okuda S, Fujita M, Itoh M*, et al.* Extracting sequence motifs and the phylogenetic features of SNARE-dependent membrane traffic. Traffic 2006; 7:1104-18.
[50] Lin PJ, Williams WP, Luu Y, Molday RS, Orlowski J*, et al.* Secretory carrier membrane proteins interact and regulate trafficking of the organellar (Na+,K+)/H+ exchanger NHE7. J Cell Sci 2005; 118:1885-97.
[51] Edeling MA, Smith C, Owen D. Life of a clathrin coat: insights from clathrin and AP structures. Nat Rev Mol Cell Biol 2006; 7:32-44.
[52] Schmid EM, McMahon HT. Integrating molecular and network biology to decode endocytosis. Nature 2007; 448:883-8.
[53] Chua CE, Lim YS, Tang BL. Rab35--a vesicular traffic-regulating small GTPase with actin modulating roles. FEBS Lett 2010; 584:1-6.
[54] Parton RG, Simons K. The multiple faces of caveolae. Nat Rev Mol Cell Biol 2007; 8:185-94.
[55] Parton RG. Caveolae--from ultrastructure to molecular mechanisms. Nat Rev Mol Cell Biol 2003; 4:162-7.

[56] Hansen CG, Nichols BJ. Molecular mechanisms of clathrin-independent endocytosis. J Cell Sci 2009; 122:1713-21.
[57] Sottile J, Chandler J. Fibronectin matrix turnover occurs through a caveolin-1-dependent process. Mol Biol Cell 2005; 16:757-68.
[58] Doherty GJ, McMahon HT. Mechanisms of endocytosis. Annu Rev Biochem 2009; 78:857-902.
[59] Blonder J, Terunuma A, Conrads TP, Chan KC, Yee C*, et al.* A proteomic characterization of the plasma membrane of human epidermis by high-throughput mass spectrometry. J Invest Dermatol 2004; 123:691-9.
[60] Godfrey RW. Human airway epithelial tight junctions. Microsc Res Tech 1997; 38:488-99.
[61] Knight D. Increased permeability of asthmatic epithelial cells to pollutants. Does this mean that they are intrinsically abnormal? Clin Exp Allergy 2002; 32:1263-5.
[62] Backhed F, Meijer L, Normark S, Richter-Dahlfors A. TLR4-dependent recognition of lipopolysaccharide by epithelial cells requires sCD14. Cell Microbiol 2002; 4:493-501.
[63] Moi P, Chan K, Asunis I, Cao A, Kan YW. Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region. Proc Natl Acad Sci U S A 1994; 91:9926-30.
[64] Li W, Kong AN. Molecular mechanisms of Nrf2-mediated antioxidant response. Mol Carcinog 2009; 48:91-104.
[65] Nguyen T, Nioi P, Pickett CB. The Nrf2-antioxidant response element signaling pathway and its activation by oxidative stress. J Biol Chem 2009; 284:13291-5.
[66] Hayes JD, Chanas SA, Henderson CJ, McMahon M, Sun C*, et al.* The Nrf2 transcription factor contributes both to the basal expression of glutathione S-transferases in mouse liver and to their induction by the chemopreventive synthetic antioxidants, butylated hydroxyanisole and ethoxyquin. Biochem Soc Trans 2000; 28:33-41.
[67] Solis WA, Dalton TP, Dieter MZ, Freshwater S, Harrer JM*, et al.* Glutamate-cysteine ligase modifier subunit: mouse Gclm gene structure and regulation by agents that cause oxidative stress. Biochem Pharmacol 2002; 63:1739-54.
[68] Miller I, Eberini I, Gianazza E. Proteomics of lung physiopathology. Proteomics 2008; 8:5053-73.
[69] Bartlett NW, Walton RP, Edwards MR, Aniscenko J, Caramori G*, et al.* Mouse models of rhinovirus-induced disease and exacerbation of allergic airway inflammation. Nat Med 2008; 14:199-204.
[70] Alom-Ruiz SP, Anilkumar N, Shah AM. Reactive oxygen species and endothelial activation. Antioxid Redox Signal 2008; 10:1089-100.
[71] Bustelo XR, Sauzeau V, Berenjeno IM. GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. Bioessays 2007; 29:356-70.
[72] Sakumura Y, Tsukada Y, Yamamoto N, Ishii S. A molecular model for axon guidance based on cross talk between rho GTPases. Biophys J 2005; 89:812-22.
[73] Weed SA, Karginov AV, Schafer DA, Weaver AM, Kinley AW*, et al.* Cortactin localization to sites of actin assembly in lamellipodia requires interactions with F-actin and the Arp2/3 complex. J Cell Biol 2000; 151:29-40.
[74] Lasonder E, Ishihama Y, Andersen JS, Vermunt AM, Pain A*, et al.* Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. Nature 2002; 419:537-42.

[75] Corbin RW, Paliy O, Yang F, Shabanowitz J, Platt M, *et al.* Toward a protein profile of Escherichia coli: comparison to its transcription profile. Proc Natl Acad Sci U S A 2003; 100:9232-7.
[76] Moll R, Divo M, Langbein L. The human keratins: biology and pathology. Histochem Cell Biol 2008; 129:705-33.
[77] Fliegauf M, Benzing T, Omran H. When cilia go bad: cilia defects and ciliopathies. Nat Rev Mol Cell Biol 2007; 8:880-93.
[78] Bingle CD, Gorr SU. Host defense in oral and airway epithelia: chromosome 20 contributes a new protein family. Int J Biochem Cell Biol 2004; 36:2144-52.
[79] Carvalho-Oliveira IM, Charro N, Aarbiou J, Buijs-Offerman RM, Wilke M, *et al.* Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009; 8:3606-16.
[80] Mark M, Ghyselinck NB, Chambon P. Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. Annu Rev Pharmacol Toxicol 2006; 46:451-80.
[81] Niederreither K, Vermot J, Schuhbaur B, Chambon P, Dolle P. Embryonic retinoic acid synthesis is required for forelimb growth and anteroposterior patterning in the mouse. Development 2002; 129:3563-74.
[82] Steiling K, Kadar AY, Bergerat A, Flanigon J, Sridhar S, *et al.* Comparison of proteomic and transcriptomic profiles in the bronchial airway epithelium of current and never smokers. PLoS One 2009; 4:e5043.
[83] Chapman AL, Morrissey BM, Vasu VT, Juarez MM, Houghton JS, *et al.* Myeloperoxidase-dependent oxidative metabolism of nitric oxide in the cystic fibrosis airway. J Cyst Fibros 2010; 9:84-92.
[84] Thomson E, Brennan S, Senthilmohan R, Gangell CL, Chapman AL, *et al.* Identifying Peroxidases and Their Oxidants in the Early Pathology of Cystic Fibrosis. Free Radic Biol Med 2010.

Chapter V

Deep Proteome Profiling of Nasal Epithelial Cells: Consequences for Impaired Respiratory Function in Cystic Fibrosis Patients

Charro N[1,2], Simões T[1,2], Faria D[5], Pacheco P[3], Lopes C[4], Azevedo P[4], Bugalho de Almeida A[4], Couto FM[5], Chan KC[2], Isaaq HJ[2], Waybright T[2], Veenstra TD[2], Blonder J[2], Penque D[1]

[1]*Laboratory of Proteomics, Dept. of Genetics, INSA Dr. Ricardo Jorge, Lisbon, Portugal*
[2]*Laboratory of Proteomics and Analytical Technologies, SAIC-Inc., NCI at Frederick, MD, USA.*
[3]*Laboratory of Molecular Biology, Dept. of Genetics, INSA Dr. Ricardo Jorge, Lisbon, Portugal.*
[4]*Clínica Universitária de Pneumologia, Hospital de Santa Maria, Faculdade de Medicina da Universidade de Lisboa.*
[5]*Dept. of Informatics, Faculdade de Ciências da Universidade de Lisboa, Portugal.*

ABSTRACT

Respiratory failure is the most life-threatening condition in CF patients and lung is the most compromised organ in this multifactorial disease caused by mutations in a gene coding for a cAMP-regulated Cl^- channel, the CFTR protein. Due to ethical and practical constrains in obtaining lung biopsies, nasal epithelial cells (NEC) can be used as surrogates of bronchial epithelia as described previously and, therefore be used in the study of CF lung disease.

Here, a wider comparative study of NEC proteome from CF patients with mild or severe lung disease in comparison with that from healthy CF-carriers or non-CF carriers was performed using a combination of cell fractionation techniques and two-dimensional chromatographic separation associated with tandem MS analysis. About 3000 proteins were confidently identified and compared across conditions. Functional characterization of the differentially expressed proteins in the NEC membrane fraction highlighted altered antigen presentation (ANXA 1 and 2, CTSG, EPX, LYZ), mitochondrial dysfunction with impaired energy production (COX4I1, COX6C, SLC25A4) and imbalanced inflammatory response (CYBA, CD63, ITGAM, ANPEP). In the NEC soluble fraction, the main regulated functions corresponded to free radical scavenging (SOD 1 and 2, CAT, GPX1), protein synthesis and folding (HSPA5, HSP90AA1, PDIA6, ERP29) and degradation of extracelular matrix (CTSB, CTTN, MMP9). Several identified proteins have been already associated with respiratory diseases (LPLUNC, LTF, PPIA, ENO1, ALDOA, HSP90AB1, S100P, CTSZ, CBR1). Altogether, these results emphasize further value of using native NEC in CF investigation by showing primary molecular networks/pathways which involve proteins that may constitute candidate biomarkers for CF lung disease.

Keywords: Cystic Fibrosis; Fractionation; Nasal Epithelial Cells; Label-free Proteomics; Pathway Analysis

1. BACKGROUND

Investigating proteome profiles and exploring altered proteins' expression as consequence of pathology is the main focus of proteomic studies. Choosing the most adequate biological sample to achieve such purpose is the first challenge to ensure that reliable results and conclusions are obtained. Monogenic disease cystic fibrosis (CF) is classically characterized by complications in the gastrointestinal tract, reproductive system with the most severe consequences occurring at the lung level [1,2]. Persistent pulmonary infections, airways obstruction, remodeling and fibrosis remain the most life-limiting condition resulting in patients' death due to lung failure. Mutations in the CF transmembrane conductance regulator (*CFTR*) gene are translated into an altered CFTR protein that grounds cause abnormal chloride transport in the respiratory epithelium and alters the mucocilliary clearance of airway secretions that become plugged and trapped creating favourable environment to opportunistic pathogens to colonize the airways [3,4]. Ideally, lung tissue would be the desirable biological sample to work with when dealing with CF's respiratory disease but collecting lung tissue for investigation studies out of the medical context can raise ethical constrains. Several studies focused their investigation into bronchoalveolar lavage fluid (BALF) [5-7], sputum [8-11], mucus [12,13] or airway surface liquid (ASL) [14,15] as means to study lung's destruction in CF by its leakage products but few dedicate it to the cellular level [16-18].

We have recently reported the most complete to date characterization of the nasal epithelia cells' (NEC) proteome (for details, see chapter IV) and made proof-of-concept about using NEC as surrogates of bronchial cell to mimic lungs' microenvironment [19-27] in investigation of respiratory diseases.

Since the discovery of the *CFTR* gene in 1989 [28,29] it has become clear that genotype differences and CFTR impairment alone do not explain the entire complex CF clinical phenotypes [30-32] and that secondary factors significantly influence the severity of lung disease [33-35]. Numerous modifier genes have been purposed and include those whose products that may improve or correct abnormal chloride transport of CFTR, modulate the cycle of infection/inflammation, regulate oxidant/antioxidant status and trigger airways' reactivity.

Using an improved strategy in sample fractionation and protein analysis, it is our goal to increase the knowledge on pathways, processes and proteins actively involved in CF's NEC with influence in lungs' disease phenotype.

2. MATERIALS AND METHODS

2.1 Demographics and Sample Collection/Preparation

Ethical approval and informed consent were obtained from participating institutions and enrolled individuals, respectively. NEC were collected by a brushing procedure previously described by our group [36] from 26 CF patients free of nasal complications and undergoing standard medical treatment, namely trobamycin and nebulized recombinant human dornase alpha, to reduce sputum's viscosity and control respiratory exacerbation. CF patients were further characterized as mild or severe according to their respiratory function based on clinical criteria and spirometry measurements [37] (Table V.1)

Table V.1: Demographics of Cystic Fibrosis (CF) patients and Carriers used for NEC collection. Disease severity was accessed by spirometry measurements and clinical evaluation according to [37]. Healthy individuals were chosen to match similar demographic conditions as the CF patients.

	Group				
	Controls	**CF Patients**			**Carriers**
		Total	**Mild**	**Severe**	
n	23	26	9	17	16
Age (mean ± SD)	(39±9) y	(27±8) y	(27±9) y	(27±8) y	(42±9) y
Gender	15 F 8 M	15 F 11 M	6 F 3 M	11 F 6 M	10 F 6 M
Genotype ΔF/ΔF (%)	NA	7 (27%)	1 (11%)	7 (41%)	NA
Genotype ΔF (%)	NA	19 (73%)	8 (89%)	10 (59%)	9 (56%)
FVC % (mean ± SD)	93 ± 11	77 ± 24	89 ± 23	71 ± 23	---
FEV_1 % (mean ± SD)	91 ± 12	60 ± 27	80 ± 24	51 ± 24	---
Colonization P. aeruginosa (%)	---	60	45	67	---
BMI (mean ± SD)	---	21.1 ± 3.5	22.9 ± 3.6	20.3 ± 3.2	---

Sixteen healthy carriers of one mutated *CFTR* allele, mostly parents of CF individuals, and 23 age- and gender-matched healthy (free of respiratory complications; FEV_1 %≥80%) individuals previously genotyped for the most common mutations present in the Portuguese population [38,39] to exclude potential asymptomatic CF-carriers/patients were used for comparison purposes. After collection, membrane and soluble subproteomes were isolated by ultracentrifugation and analysed by label-free

2D-LC-MS/MS as previously described by us (chapter IV) to obtain proteins' identification and quantitative information on unique and total peptides sequenced. A list of proteins with the correspondent unique and total peptide counts (UPC and TPC, respectively) for each specimen in each condition analysed was obtained.

2.2 Data Normalization and Relative Protein Quantification using a Label-Free Approach

2.2.1 Relative Quantification by Total Signal (TS)

From the total list obtained, proteins were first filtered by TPC $\geq$ 2 and normalized by total signal (TS) approach according to Carvalho *et al.* [40]. Briefly, for each condition, TS is calculated by the sum of all individual protein's TPC (iTPC) and used as denominator in the fraction were the numerator is the TPC for a particular protein (iTPC/TS). This process is repeated along the list of proteins, therefore normalizing the counts for each protein in the considered condition. Relative quantification is achieved by comparing these normalized TPC (normTPC) for the same protein across the different conditions analysed and proteins evidencing a fold change of ± 1.5 were considered as differentially expressed and further explored in terms of biological and functional relevance.

2.2.2 Relative Quantification by exponentially modified Protein Abundance Index (emPAI)

Relative quantification was also estimated by exponentially modified protein abundance index (emPAI) previously described by [41] and [42] as there is general assumption that the number of peptides sequenced *per* protein correlates with the protein's content in the mixture. This index is a handy and easily computed indicator that can be used to produce protein expression data from any LC-MS-based strategy, especially in cases where isotope-based approaches cannot be applied. Obtained emPAI values of proteins were then compared across samples. Proteins displaying at least a ± 1.5 fold difference in abundance were considered as differentially expressed.

2.2.3 Data Analysis

As similarly performed in the full characterization of the NEC proteome (chapter IV), hydrophobic character (GRAVY) and putative transmembrane domains (TMD) were calculated by Gene Infinity tools (http://www.geneinfinity.org/sms_proteingravy.html) and Centre for Biological Sequence TMHMM server (http://www.cbs.dtu.dk/services/TMHMM), respectively.

Interpretation of the differentially expressed proteins was performed by Ingenuity Pathway Analysis (IPA) knowledgebase, online tool Protein Information Knowledge Extractor (PIKE) and Gene Ontology (GO) to integrate those proteins into pathways with biological significance while exploring their individual function and localization.

2.2.4 Western Blot

30 μg of protein from total protein extract of NEC from the groups under analysis were separated on 4-12 % (w/v) polyacrylamide gels, transferred to nitrocellulose membranes (Schleicher&Schuell) and analyzed by western blot, each triplicate analyzed in independent gels. Membranes were probed with mouse monoclonal anti-GPX1 (Abcam 68360) (10μg/ml) in PBS supplemented with 5 % (w/v) fat free milk for 2 h at ambient temperature and developed using chemiluminescence-ECL (Pierce). After development, membranes were washed with stripping buffer [1.5% glycine (w/v), 0.1% SDS (w/v), 1% Tween-20 (v/v), pH 2.2] 5 times for 10 min, washed with PBS 6 times for 5 min before blocking with PBS supplemented with 5 % (w/v) fat free milk overnight. At this point, the membrane is ready to be incubated with another antibody. Following the same protocol, membranes were probe with anti-ATP5A1 (Abcam 14748) (0.2μg/ml). The abundance of selected proteins was calculated from densitometry of immunoblots using Progenesis PG200v2006.

3. RESULTS AND DISCUSSION

There has been some controversy over the use of NEC as a surrogate model of lower airways portion in lung pathophysiology studies. Some authors claim that NEC are not representative of bronchia and bronchioles while others argue that there seem to be more similarities than differences between upper and lower airways. However, the concept of 'united airways' where changes in the first might seriously compromise the function and structure of the latest has been generally accepted. We have previously characterized NEC's proteome by a broad high-throughput technology highlighting proteins, functions and pathways active in this epithelium while establishing a proteome portrait that evidence the properties of lower airways. This is of particular biological relevance for the study of pulmonary diseases, such as Cystic Fibrosis (CF), a monogenic disease compromising the function of several organs but where respiratory complications and lung failure are the most life-threatening conditions. Two objectives drove the design of the present work: to identify differences in protein abundances by comparing (1) healthy non-CF, CF-carriers and CF patients (genotype-based analysis) and, among those, (2) patients with mild or severe lung disease categorized based on FEV_1 % predicted accordingly to Johnson *et al.* [37] (respiratory-based analysis). Previous studies in primary respiratory cells [43], mice [44,45], BALF and sputum [5,7,8,10,11,46] and also in NEC [16-18], the same biological sample used here, have provided comparisons between CF and non-CF. Our work, however, provides the deepest and most complete to date proteome comparative analysis of these cells indicating the imbalanced proteinaceous environment and possible consequences in welfare of CF patients. As so, NEC are used here as a model to investigate the molecular determinants in the development and/or progression of CF respiratory disease.

3.1 General considerations on the differentially expressed proteins identified in NEC from CF patients

We fractionated NEC into membrane and soluble/cytoplasmic subproteomes by ultracentrifugation and incubation in specific buffers to simplify and enhance the identification of low abundant proteins. The false discovery rate (FDR) calculated by searching the data against a decoy database was 4.5%.

Hydrophobicity and predicted number of TMD of the differentially expressed proteins identified here are in accordance with previous results, showing an enrichment of highly hydrophobic specimens with more putative TMD in the membrane fraction, whereas proteins with lower hydrophobicities and absent (or reduced, in a minority of cases) TMD are mainly allocated to the soluble portion. The overlap fraction between the two subproteomes encompasses an intermediary situation. Pre-cytological analysis evidenced that the majority of cells recovered are epithelial, fact corroborated by the identification of several epithelial markers such as KRT 5, 8, 13, 18 and MUC-5AC, all being up regulated in patients' samples.

3.2 Relative quantification strategies

Label free strategies are based on the assumption that the abundance of a protein in a mixture is proportional to the number of its peptides identified. However, inter runs variability occurs and different methods to correct it have emerged allowing the results to be comparable. Here, we used two parallel strategies: normalization by total signal (TS) (Tables V.1.1SD and V.1.2SD) and exponentially modified protein abundance index (emPAI) (Tables V.2.1SD and V.2.2SD). The number of differentially expressed proteins obtained by each method is significantly different: TS relative quantification retrieved 512 differentially expressed proteins regardless the subproteome analysed while emPAI quantification showed a total of 1291 proteins (Table V.1SD and V.2SD). It is noteworthy to say that these numbers depend largely on the initial lists that are used for standardization. The TS normalization implies an *a priori* filtering on proteins identified by at least TPC $\geq$2 to sustain that identification is reliable, whereas emPAI normalization uses all proteins identified in the dataset (UPC $\geq$1) assuming that, if a protein is detected even by a single non-redundant peptide, than it must be present in the mixture. Yet, when emPAI normalization is performed on the data prefiltered with the TS criteria, the numbers converge to ~500 proteins and the expression tendencies are maintained in both methodologies.

3.3 Expression patterns of the differentially expressed proteins

Interpretation of differentially expressed proteins was facilitated by construction of proteome profiles where tendencies of expression are shown (Figures 1 and 4) among the subcompartments isolated during sample preparation (membrane and soluble subproteomes). Ultimately, our goal is to provide protein patterns where CF patients can

be effectively distinguished from healthy individuals as well as other patients sharing CF-related symptoms.

3.3.1 Proteins from the membrane fraction

Concerning the specimens identified in the membrane compartment, 260 proteins were found as differentially expressed by TS method, 242 of them (93%) in common with emPAI standardization.

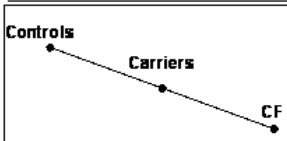
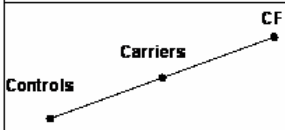
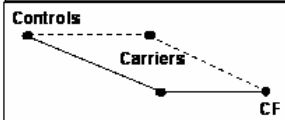
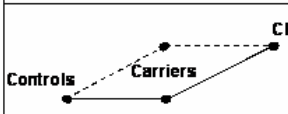
The expression patterns created allowed the recognition of proteins consistently up or down regulated in the groups under comparison (Figure 1A and 1B). We observed EPX, CD44 antigen, CTNND1, KRT8, PLEC1 and JUP with consistently with increased abundances from healthy to CF carriers and then to patients. Except for JUP, higher levels of these proteins were also observed in CF patients with severe lung disease when compared with mild ones.

3.3.1.1 Proteins involved in cytoskeleton and tissue remodeling

Catenin δ-1 (CTNND1) belongs to a large multiprotein cell-cell adhesion complex that associates and regulates cellular adhesion of C- and E-cadherins, important to build-up a structured epithelium [47,48]. A study performed in human and rat cell lines showed that CTNND1 is more expressed in epithelial cells than in fibroblast (mesenchymal origin), although this expression was not cell-type restricted [49]. Work from Castillo *et al.* showed strong immunostaining for CTNND1 in lung tumours, making it a candidate for anticancer therapies [50]. Also, CTNND1 is induced by wounding [51], circumstance observed at the respiratory level in CF. Epithelial origin of CTNND1 and its involvement in lung pathology provides good correlation with CF lung disease.

Cytoskeleton proteins are responsible for the scaffolding of proper cellular arrangements. Here, plectin-1 (PLEC1) and cytokeratin 8 (KRT8) have increased abundances in CF carriers and even higher levels in patients. Plectin-1 interacts with actin, cross-linking and stabilizing cytoskeletal intermediate filaments network but also regulating their dynamics [52,53]. Actins and actin-associated proteins (ACTB, ACTA2, PLXNB2) were identified in both membrane and soluble fraction, being more abundant in diseased. These proteins are fundamental in epithelial cells migration events but also in inflammation and immune response.

mNEC – Mutational-based Analysis

	SAMM50; CD97; CD59; SLC25A24; MTCH1; TMEM33; TMEM30A; ERMP1; CLDN1; PTPLAD1
	EPX; CTNND1; CD44; KRT8; PLEC1; JUP
	RPS2; RPS13; RPS16; RPS25; RPL7; RPL19; TMED9; TOMM40; ALDH3A2; ETFDH; BPI; NUP210; UQCRC2; CD9; MTCH2; SMPD2; PEX11B; BRP44; RAP1A; SLC25A5 **DNACJ11; TMED4; TMEM109; DES; A2M; PLP2; ABCD3; OAS3; LYZ; LAMP1; LMAN1; UQCRC1; CYBA; TM9SF2; CAMP; TLR3; LGALS3; F11R; ABCC4; RPL10L; CTSG**
	PPIB; PLXNB2; DLAT; NDUFB10; HNRNPU; ACTB; ACTA2; FTGFRN; DHRS7B; SCARB2; KRT7; ANXA2; STT3B; B3GAT3; DHRS7; RPS3; RPS18; SFXN3; TUFM; TUBA1A; TUBB1; TUBBAC; GALNT4; HK1; ATP2A3; ATP5A1; ATP12A; RAB14; HLA-A; MUC-5AC; TRPM4; CPT1A; HIST1H2BB **ATP13A1; CYP20A1; SLC16A3; IGHG1; KRT1; KRT4; KRT5; KRT13; KRT18; ATP5L; HIST1H2AA**

mNEC – Respiratory Function-based Analysis

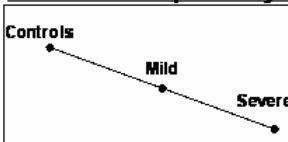
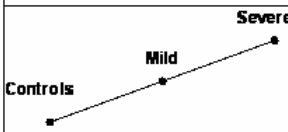
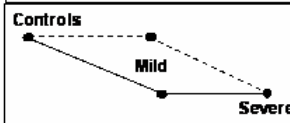
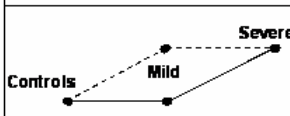
	PLP2; CYP2S1; TM9SF2; LYPD2; LYZ; SMPD2; CD82; PPIB; TMEM33
	IMMT; COX6C; HLA-B; RPL3; RPL6; LETM1; KRT7; ACTA2; ATP5A1; HIST1H2BB
	MAOA; FUNDC2; IGHA1; ITGA3; OAS3; TM9SF1; BPI; RPL19; RPS25; RPS27A; RPS9; METTL7A; SYNGR1; STT3A; COX5A; TSGA10; PRDX1; PTGES; MGST3; ATP2A1; ALDH3B1; PGRMC2; YIFI1A; AQP5; UGT2A1; CEACAM1; RDH14 **TMED1; TMEM109; RPS17; DNAJC11; CLDN1; DAD1; COX5B; MTCH1; ITGB2; LTF; BASP; CD59; CD97; RPL7; RPL10L; OCIAD2; DHODH; MT-ND4; UQCRC1; CTSG; TLR3; SLC44A2; SLC25A24; SLC25A11; SLC35B2; DES; ETFDH; CDIPT**
	KRT2; KRT4; KRT5; KRT6A; KRT8; KRT13; KRT14; KRT18; VAMP8; NDUFA9; HP1BP3; EPX; MUC-5AC; CTNND1; TUBA1A; PPOX; ANXA2; EEF1A1; CD44; ATP12A; PLEC1; HK1; RPS24; RPN1; SAMHD1; HSPD1; SUN1; RPL7A; ACTB; ITPR3; DHRS7B; GALNT4; TMCO1 **HBB; RPS2; RPS18; LRRC59; ATP5F1; SLC25A13; MT-ND1; ANXA1; ESYT1; GOLM1; B3GNT3; DHRS7; ANPEP; SFXN3; LMAN2**

Figure V.1: Proteome profiles of the differentially expressed proteins in the membrane fraction between healthy non-CF, CF carriers and CF patients (**1A**) and between healthy non-CF and CF patients with Mild or Severe CF lung disease (**1B**).

Tubulin's isoforms (TUBB1, TUBBAC, TUBA1A) are present in high abundance in NEC as previously determined and are another group of cytoskeleton proteins with higher abundance in CF patients with similar levels in healthy CF-carriers and non-CF carriers. Their importance is not limited to microtubules constituents; additional roles in airway epithelium remodeling at the tall columnar cells' level have been purposed [54] as well as taking part in SARS [55] and non-small cell lung carcinoma [56,57].

Cytokeratins belong to a protein's family with high abundance in NEC. In general, keratins are part of the innate machinery responsible for the stability from a single cell to a fully organized and functional tissue [58]. Their patterns of expression can be regarded as specific markers for the epithelial differentiation status in the respiratory tract. Among the cytokeratins identified, KRT 5, 8 and 13 are typical of the respiratory epithelium [59]. KRT8 has already been associated with fibrosis of the exocrine pancreas [60], the most well-established phenotype-genotype correlation in CF [61,62]. Its presence mainly at the lower airways level (bronchial, alveolar epithelium and pleural mesothelium), along with KRT7 [63], makes it a good target for investigation of its role in remodeling and wound repair processes.

KRT18 is the most ubiquitously expressed keratin in the respiratory tract [63] and plays a key part in delivering mutated CFTR to the plasma membrane [64,65] and, together with KRT8, associates with IL-6 to trigger a neutrophil-mediated response to aggression from bacteria and viruses [66,67]. In CF, the physiochemical changes in the airways secretions that impair innate defence result in chronic endobronchial infection and in a robust neutrophilic inflammatory response [68]. Higher levels of autoantibodies against KRT8 and KRT18 were also found in CF's sera when compared with control (*unpublished work*). In brief, keratins have an undoubtedly role in filament organization crucial for the maintenance of cellular assembly and structure with additional functions in inflammatory and immunity events.

3.3.1.2 Mucus viscosity and water transport

Both CD44 and MMP9, a major component of the extracellular matrix metalloproteinases (MMP) family with correlation with lung disease (discussed below), present crescent abundances from healthy non-CF to CF carriers and from those to CF patients with no significant differences between mild and severe patients. CD44 antigen is a cell-surface glycoprotein involved in cell-cell interactions, cell adhesion and migration in inflamed tissue that was shown to be metastatic determinant of rat pancreatic tumour. In human, increased CD44 has been associated with aggressive stages of various cancers and inflammtory response due to its ligand activity with hyaluronic acid and MMP [69,70]. Hyaluronic acid is a component of the extracellular matrix with viscoelastic and hygroscopic capacities acting as a scaffold, entrapping large amounts of water and ions to provide tissues with hydration and turgescence [71]. Overexpression of CD44 and its consequent binding to hyaluronan might be the result

of an adaptative strategy to increase lubricant properties in CF airways. In accordance with this, MUC5AC was found increased in abundance in CF carriers and patients when compared with controls and similar levels were found in mild and severe patients. MUC5AC is a gel-forming glycoprotein highly expressed in the surface mucosal cells of respiratory and gastric epithelia that protects the mucosa from infection and chemical damage by binding to inhaled microrganisms and particules that are subsequently removed by the mucocilary system [72-75]. Hypersecretion and retention of dehydrated mucus contribute to poor transportability of allergens plugging CF's airways and compromising the respiratory function [76]. Regulating water absortion and secretion can contribute to ameliorate/compromise the airways' obstruction seem in CF. Water transport is driven by osmosis, determined by NaCl movement and occurs through the lipid bylayer and aquaporin proteins [77]. AQP5 forms a water-specific channel implicated in the generation of saliva, tears and pulmonary secretions, and was identified less abundant in severe CF patients than in mild or healthy individuals. This protein was found to be widely distributed throughout the surface epithelia in mouse respiratory system mainly at the luminal side of epithelial cells from the nasal cavity to intrapulmonary bronchioles in contrast to previously reported restriction to type I alveolar cells [78]. Its critical role in water handling such as the maintenaince of airway surface liquid (ASL) and clearance of alveolar fluid have implicated it in airway inflammation and mucus clearance in the development of asthma [79]. These authors also observed significantly lower expression of MUC5AC and MUC5B in lungs of AQP5-knockout mice. Others have shown that depletion in AQP5 resulted in increased MUC5AC secretion [80]. Our results corroborate this latter effect with higher levels of MUC5AC when AQP5 presented lower abundance. Regulatory mechanisms between aquaporins and mucins may highlight new strategies for novel antihypersecretory drugs in airway diseases.

3.3.1.3 Antibacterial enviroment in CF's lungs

One of the main functions of airway epithelium is to sense and actively respond to alterations in its surrondings [81]. We were able to identify differential regulation in several proteins involved in antibacterial response of lungs. All Lyz, CAMP, LTF and BPI have primarily a bacteriolytic function, actively involved in deplete the airways from opportunistic pathogens that colonize CF's lungs [3,4,82,83]. In fact, BPI has antibacterial activity against *Pseudomonas aeruginosa* but its inhibited by its bacterium

LPS [84,85] which explains its lower abundance in CF group. CTSG belongs to the peptidase module compromising 22 neutrophil-derived proteins implicated in tissue injury characteristic of CF lung disease [86-88] and also presents antibacterial activity against *P.aeruginosa* and inhibition by LPS [85]. Although less abundant in CF carriers and patients with equivalent levels when respiratory function is compared, CTGS inhibits IL-8, the major neutrophil chemoattractant peptide and previously reported marker of inflammation in CF [89,90]. Reduction in CTSG conducts to inneficient inhibiton of this proinflammatory pathway, perpectuating self-cycle of inflammation in CF's airways.

Two members of the major histocompatibility complex (MHC) (HLA-A and HLA-B) and MHC binders (IGHG1, IGHA1 and ITGA3) were found differentially expressed. These components may serve both to defend against local infection and to prevent access of foreign antigens to the general immunologic system. Circulating autoantibodies against IGHG1 were described by Pedersen and colleagues [9] in sputum of CF patients and have previously being described in autoimmune [91] and chronic inflammatory disorders [92].

Members of the Toll-like receptors (TLR) family are expressed on a large number of immune and epithelial cells [93] with impact on activation and maintenance of the innate immune response to microbial pathogens [94]. Here, we identified TLR3 with reduced abundance in CF carriers and patients regardless their lung's function severity in comparison with non-CF. Its influence in asthma and COPD has been demonstrated [95,96] with increased mucus production, IL-13 and IL-5 expression and eosinophils' recruitment in the airways of TLR3-knockout mice infected with respiratory syncytial virus [97]. Although TLR3 may not be required for viral clearance, it is necessary to maintain a proper immune environment in the lung to avoid developing pathologic symptoms of disease. This might be responsible for changes in EPX, an eosinophil peroxidase showing increased abundance from healthy to CF carriers and from those to patients. Mild and severe patients present identical levels for EPX though higher than in non-CF individuals. Eosinophils accumulate at sites of parasite invasion with beneficial roles to eliminate tissue-invasive specimens [98] and in the pathogenesis of various common human allergic diseases.

In a EPX-knockout mice model of asthma, levels for bromotyrosine, a sensitive marker for EPX oxidation [99], were undetectable in their lungs [100] and found 100-fold greater in BALF of severe asthmatics than those of non-asthmatics [101]. Unlike

neutrophils that endocytose pathogens and deposite their granule contents intracellularly into target phagossomes, activated eosinophils exocytose their granules onto the surface of adherent parasite targets [98,102-104] generating high amounts of superoxide anion (O^{2-}) and hydrogen peroxide (H_2O_2) [105]. EPX's functional role in host defence and disease pathology was demonstrated in pneumocytes, causing necrotic cell blebbing of lung epithelial cells [106,107]. Although eosinophils-specific granule proteins are markedly citotoxic for parasites, its oxidant products might interact with the surronding molecules causing tissue instability.

3.3.1.4 Impairment in mitochondrias' function and energy production

Respiratory burst denotes a rapid and increased release of ROS from immune cells due to presentation with foreign entities mostly due to mitochondria's acitivity. In our work, we observed several mithocondrial proteins differentially expressed, especially at the respiratory chain level: high abundance of NDUFB10, ATP5A1 and ATP5L, CYP20A1 and CPT1A, and decrease in SAMM50, UQCRC1 and UQCRC2, DNAJC11, CYBA, ETFDH and TOMM40. When comparing patients with mild or severe CF lung disease, reduction in MAOA, COX5A and COX5B, UGT2A1, MT-ND4 and UQCRC1, and increase in COX6C, ATP5A1 and ATP5F1, NDUFA9, MT-ND1 and PPOX is observed.. The drawback of high potential energetic compounds' generation (mainly ATP) by mitochondria through the respiratory chain and oxidative phosphorylation is the generation of ROS. Altered expression of the several components described here might compromise vital functions for cellular maintenance. Our results suggest altered function of almost all the complexes of the mitochondrial respiratory chain.

Subunits of complex I or NADH dehydrogenase (ETFDH, NDUFA9, NDUFB10, MT-ND1 and MT-ND4) showed impaired expression in CF: the protein involved in the initial electronic transfer step from NADH to flavoprotein, ETFDH, is lowered in CF while NDUFA9, NDUFB10 and MT-ND1 are increased.

Two subunits of complex III, UQCRC1 and UQCRC2, and two others from complex IV, COX5A and COX5B, were also identified as leess abundant in CF while COX6C showed increasing abundance from healthy to mild CF and from those to severe lung disease patients. This latter component was identified relevant in the molecular mechanisms leading to loss in muscle force during acute exaceration in COPD patients

[108]. Members of the complex V of the respiratory chain (ATP5A1, ATP5F1, ATP5L) are responsible for producing ATP from ADP in the presence of a proton gradient across the membrane. Downregulation of ATP5A1 was already associated with renal cell carcinomas [109] and development of tumours with microsatellite instability [110].

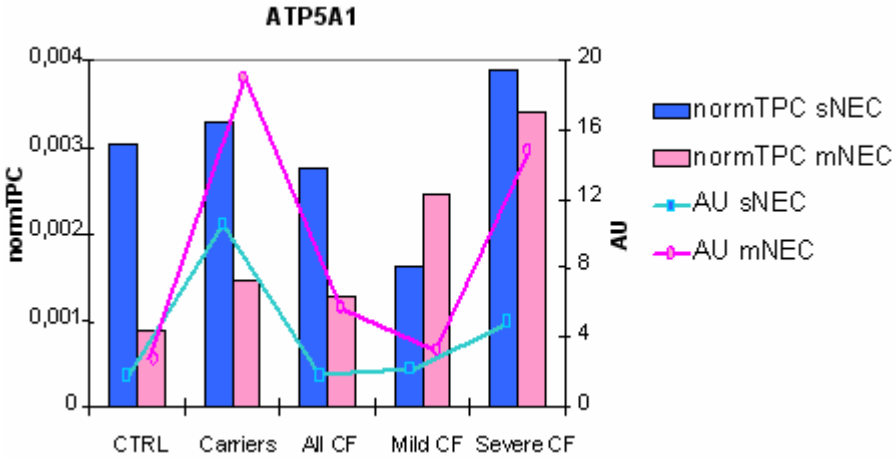


Figure V.2: Correlation between the normalized total peptide count (normTPC) and intensity band volume for ATP5A1 protein in a LC-MS/MS and Western Blot experiment respectively. (Blue bars: normTPC for ATP5A1 in the soluble fraction of NEC; Pink bars: normTPC for ATP5A1 in the membrane fraction of NEC; Blue line: arbibtraty units of the intensity of the band obtained from WB for ATP5A1 in the soluble fraction; Pink line: arbibtraty units of the intensity of the band obtained from WB for ATP5A1 in the membrane fraction)

Previous work from our group also identified ATP5A1 as downregulated in wild type and F508del-CFTR mice in response to naphthalene challenge [44]. Our results however propose an increase of complex's V activity somehow overcoming imbalanced energy production or decoupling of the various upstream complexes of the respiratory chain. Mitochondria dysfunction and uncouplig of the respiratory chain seem therefore actively involved in CF's lung disease. Imbalance in ATP production with consequent implications in cellular homeostasis may be one of the mechanisms implicated in deterioration of lung fuction.

Cytochrome P450 monooxygenases (CYP) play an important role in the defence against inhaled toxicants. In our study, we identified CYP2S1 as consistently decreased from non-CF to mild to severe patients. CYP2S1 participates in extrahepatic xenobiotics metabolism, namely naphthalene. The highest expression levels have been observed in epithelial cells from tissues frequently exposed to xenobiotics, like the respiratory or

gastrointestinal tracts [111], being highly present in BALF of smokers [112] though implicated in the activation of anticancer drugs [113]. This enzyme was shown to be induced by retinoic acid [114], an important regulator of of organ development, including the lung [115]. Our results and others have proposed that retinoic acid metabolites might be potentially used in CF treatment in a mice model [116,117]. Another evidence indicating altered xenobiotics metabolism is the reduced expression of UGT2A1 in severe CF patients compared with mild. UGT2A1 participates in elimination of potentially toxic xenobiotics and endogenous compounds in the airways whose levels are repressed in BALF of smokers [112].

3.3.1.5 Functional Analysis

Sensing the environment, reading a message and translate it into the cell characterizes the epithelium as an interface between extracelular and intracellular world. Representative networks functions and pathways on which differentially expressed proteins participate in are in accordance with such functions and represented in Tables V.3SD, V.4SD and Figures V.1SD and V.2.SD. Antigen presentation pathway is one of the means to achieve that purpose and was identified as significantly altered due to CF by pathway analysis. Several constituents of the MHC class I and II (HLA-DRB3, HLA-A, HLA-DRA, HLA-B), antigen transporter TAP2 and binders IGHG1, CD59 and CD166 were identified. Susceptibility for persistent infection of CF patients' lungs makes them more exposed to pathogen and bacterial products that trigger enhanced/inefficient response to these challenges.

Maintaining a structured and functional epithelium depends largely on interactions and adhesion to the surrounding cells. Ingenuity retrieved cellular assembly, organization and function and cell-to-cell signalling and interaction as overrepresented networks in which differentially expressed proteins in membrane fraction participate in (Figure 3A). In particular, caveolar-mediated endocytosis appears as a pathway with several of the differentially expressed proteins allocated here including members of the integrins family (ITGA3, ITGB2, ITGAM), GTPase Rab family (RAB5B, RAB5C) among others (HLA-A, HLA-B, ALB, ACTB, ACTA2, EGFR). Integrins and members of the Rab family appear to be very important in caveolar dynamics [118] and to cell adhesion phenomena [119] aiming to maintain NEC's integrity and functionality.
Membrane fraction is enriched in both plasma and microssomal membrane proteins. Therefore, organelle dynamics is also among the functions overrepresented here.

Mitochondrial and endoplasmic reticulum membrane proteins account for the main bulk of proteins with altered abundance. As so, altered mechanisms in the membrane fraction of CF'S NEC are mitochondrial dysfunction or oxidative phosphorylation. As previously postulated here by us and previously by ithers, NEC might have high demands for energy production/consumption as first line defence and barrier between environment and organism's milieu [120,121]. In CF, chronic inflammation of the lungs leads to imbalanced tissue remodeling and energetic metabolism might be a requisite to overcome this event. In fact, energy production is among the top score functions retrieved by pathway analysis using Ingenuity with several proteins contributing to that (Tables V.3SD and V.4SD and Figure 3B).

Closely related to mitochondrias' dysfunction is oxidative stress response and xenobiotics metabolism by cytochrome P450 (CYP450), other two processes highly altered as consequence of CF. Several other specimens identified in the membrane portion provide clear evidence of deregulated oxidant/antioxidant status and inability to effectively metabolize and excrete harmful compounds, emphasizing the need of combined strategies to fulfil a role. Detoxifying enzymes/proteins or master regulation of the NFR-2-mediated antioxidant response are among key intervenients in those.

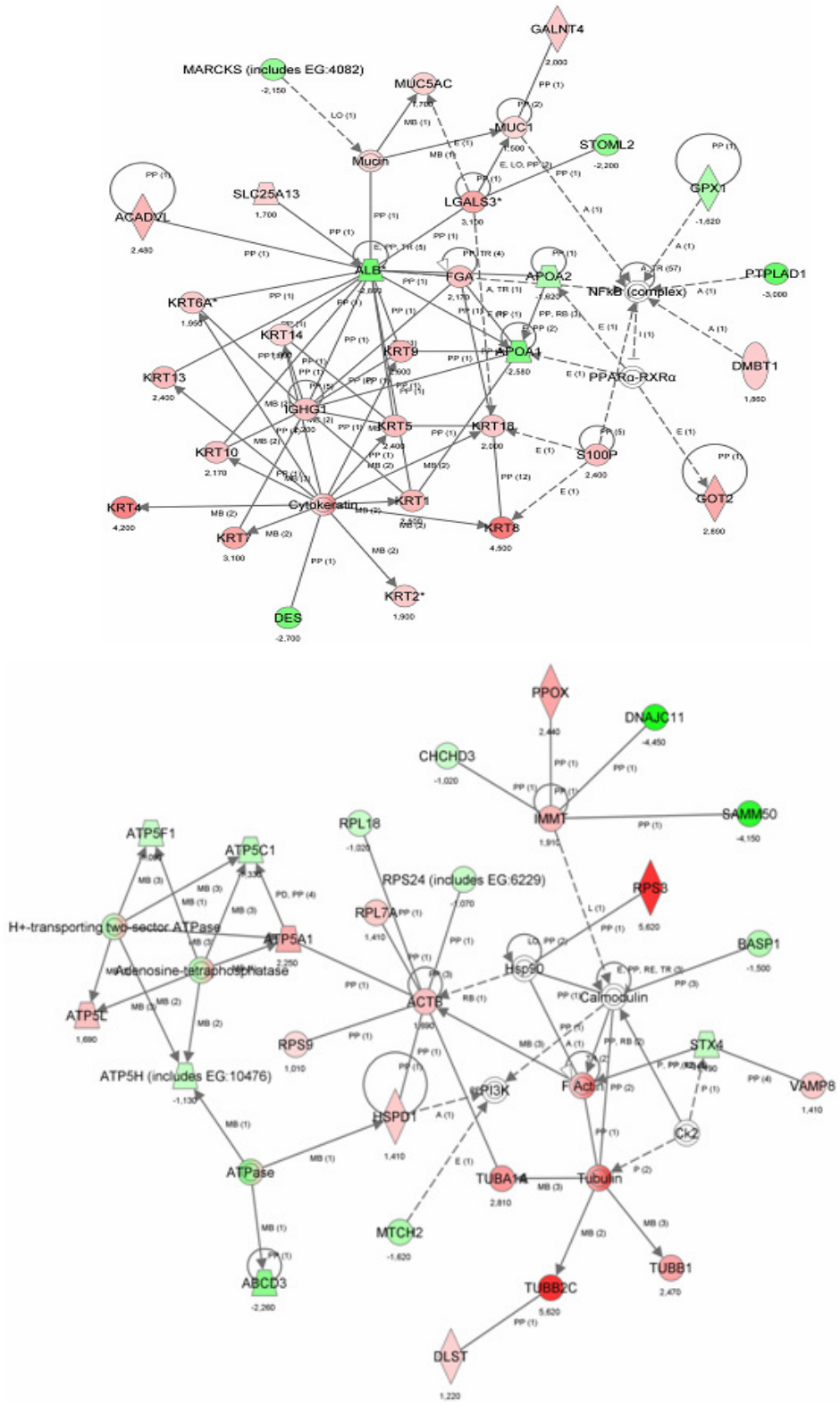


Figure V.3: Differentially expressed proteins involved in cellular assembly and organization (**3A**) and energy production and mitochondria dysfunction (**3B**), according to IPA's knowledgebase.

3.3.2 Proteins from the soluble fraction

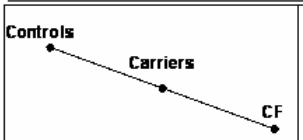
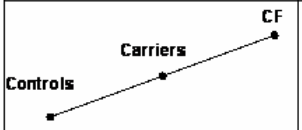
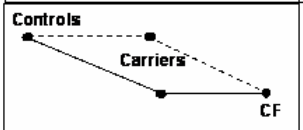
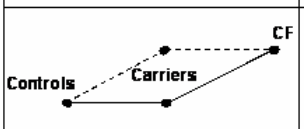
Among all the proteins identified in the soluble portion, a total of 282 differentially expressed proteins were identified by TS normalization. Of those, 239 (85%) overlap with emPAI calculations.

We were able to observe proteins progressively over expressed when comparing healthy individuals with CF carriers and patients (Figure 4).

Protein S100P is a member of small Ca^{2+}-binding proteins [122] postulated to be strongly significant in the context of different cancers [123]. Although never previously associated with CF, over expression of S100P was observed in metastasizing lung tumours with poor patient survival [124-126] and increased angiogenesis [127]. Its primary mechanism of action is through the activation of cell surface receptor of advanced glycation end products (RAGE) by stimulation of mitogen-activated protein kinase (MAPK) and NFκB [128] leading to a proinflammatory phenotype. Recurrent infections and a chronic state of inflammation characterize the lungs of CF patients [129] with NFκB pathway being actively involved in the CF's phenotype [130]. Elevated levels of S100P probably potentate persistent inflammation in CF's NEC as determined here.

MMP9 acts as an endopeptidase capable of degrading all kinds of extracelular matrix proteins, plays major roles in apoptosis, host defence and tissue remodeling, particularly at the airways level [131], contributing to airway wall thickening and inflammation [132]. Its involvement in the development of COPD and asthma, pathologies highly similar to CF, and even in lung cancer have been extensively reported [133-135]. Gharib *et al.* (2009) also showed increased levels of this metalloproteinase in BALF of CF subjects [7]. MMP9 is also involved in increasing the IL-8 proinflammatory cytokine, a signalling pathway characteristically increased in CF [136-138]. Although somehow expectable, we showed progressively elevated levels of MMP9 in NEC of CF carriers and patients that might contribute significantly to continuous lung damage. Similar regulation of the IL-8 pathway is in accordance with the results described in the membrane fraction.

sNEC – Mutational-based Analysis

	APOA1
	S100P; MMP9; RPL29; ALDH18A1; SUCLA2; CLTA; SFRS2B; CTSS; FGB; GOT2; IGHA2; PYGB; HOXA4; SLC9A3R1; SUPV3L1
	STIP1; RAN; HSPA4; SSB; HNRPAL1L2; FLNA; CES1; P4HB; CSDA; RPSA; ILF3; SERPINB4; KHSRP; GSTP1; LCP1; HNRNPD; ADH7; MARCKS **HBA1; GDI2; HADHB; BCAS1; TBCA; APOA2; UBE2N; CTTN; S100A11**
	ACTN4; PGAM1; LDHA; RPLP2; EEF1G; RPS14; AP1B1; GOT1; RBM3; ASAH1; ACTN1; KRT2; KRT10; SFN; TPM1; ILF2; HNRNPCL1; ARHGDIB; CALR; HIST1H4A; LAP3; IDH3A; CAPZB; ALDH7A1; ECH1; SPTAN1; RPS18; ACADVL; NUCB2; RBMX; SCP2; AKR1C1; DMBT1; CKMT1A; CAPZA1; ACTB; ALB; PFN2; HIST1H1A; TMPO; ACAT **RRBP1; KRT6A; KRT75; HNRNPA2B1; IDH1; EZR; NACA; PARK7; LMNB1; SUMO2; HMGCL; HBD; HLA-A; ANXA4; HSD17B8; AK3**

sNEC – Respiratory Function-based Analysis

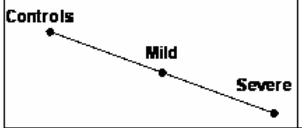
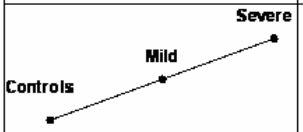
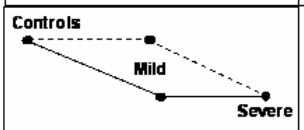
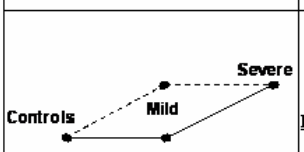
	GD12; CSDA; HBA1
	IDH3A; ANXA4; AK3; CTSS; LGALS3; AHCY; IGHA2; ECH1
	FLNA; FLNB; PPL; HNRNPA3; HNRPDL; TPI1; KRT17; CALM; PGK1; S100A4; OTUB1; VCL; EIF4B **BLVRB; HADHB; STIP1; IQGAP1; CTSZ; TBCA; RAN; SERPINB4; HSPA4; SH3BGRL3; EPB41L1; HNRNPA1L2; KHSRR; RPSA; ILF3; ANXA7; SKP1; HNRNPA0; RAVER1; TRIM29; LPLUNC**
	HMGCL; HBD; KRT2; KRT5; KRT6A; ARHGDIB; AKR1C1; DMBT1; WARS; HSPD1; LDHA; HNRNPA2B1; RBMX; NPEPPS; HNRNPUL2; S100P; GOT2; TMPO; FGB; SUPV3L1; SPTAN1; HOXA4; EZR; MMP9; ALDH18A1; SUCLA2; CALR; PFN2; HIST1H1A; SFN; HNRNPCL1; ACADVL; NUCB2 **RPL22; FGA; VHCH1; HADHA; PFN1; ALDH2; UBA1; ERP29; TUBB2C; CNDD2; PIGR; AKR1A1; LMNB1; SUMO2;HLA-A;HSD17B8; NANS; RPLP2; EEF1G; RPS14; AP1B1; RPL29; TKT; CAPZB; ALDH7A1; ASAH1; CKMT1A; CAPZA1; ALB; SLC9A3R1; CAPS; WDR1; ACAT1; SF3B1;ILF2; PYGB**

Figure V.4: Proteome profiles of the differentially expressed proteins in the soluble fraction between healthy non-CF, CF carriers and CF patients (**4A**) and between healthy non-CF and CF patients with Mild or Severe CF lung disease (**4B**).

Similarly, CTSS is an alveolar macrophage proteolytic enzyme that has been implicated in remodeling of extracelular matrix, alveolar destruction that characterizes emphysema [139] and promotion of cilia motility from conducting airway cells in the lung [140]. Elevated levels of CTSS were found in lung tumours when compared with control parenchyma from the same subjects, although serum levels did not significantly change, revealing possible roles in tumour growth and progression [141]. Its localization in lymphatic tissues and antigen-presenting cells however strongly suggest

additional roles of CTSS in modulating the immune response with potential therapeutic use in asthma, hypersensitive pneumonitis and autoimmune disorders [140,142]. In CF, the autoimmune component has been postulated [9,143-146] while impaired and excessive tissue damage and repair occurs especially at lung level. Increased levels of CTSS in carriers and CF patients, with higher expression in severe than in mild lung disease, were consistent with the broad phenotypes evidenced by these patients. Other members of the cathepsins family were identified here such as CTSB and CTSZ. CTSB is an ubiquitously expressed protein although moderate levels have been observed in bronchus and nasopharynx at the respiratory epithelial cells level and in macrophages in lung [147]. Levels for this protein were similar when comparing CF patients with healthy individuals although a small increase was observed in CF carriers. Two different studies reported CTSB as down regulated in CF [16,17] while its activity was reported to be elevated in sputum of inflamed CF-airways [148] and CF serum of patients [149]. The levels present here might reflect a controlled and non-acute phase of inflammation of the enrolled individuals. On the other hand, CTSZ presented lower levels in CF than in healthy with similar levels when respiratory function was concerned. The lowest levels were detected in CF carriers. Both CTSB and CTSZ have been implicated in cancer pathogenesis showing synergistic anti-tumour effects in double knockout mice with a 70% and 80% reduction in metastases number and size, respectively [150]. Association with SARS has also been reported.

NHE-RF1 is another protein elevated from healthy to CF carriers and from those to patients. Severe patients also presented higher levels of this protein when compared with mild ones. NHE-RF1 was the first identified CFTR-binding protein and association with CF has been extensively reported since it [151]. NHE-RF1 is a scaffold protein that connects plasma membrane proteins with members of the ezrin/moesin/radixin (ERM) family aiding their linkage to the actin cytoskeleton and regulating their surface expression. Especially, the interaction of NHE-RF1 and CFTR plays a central role in regulating the CFTR trafficking to the apical membrane and stabilizing it on airway epithelial cells [152,153], making it essential for both normal and mutant CFTR function [154]. It was demonstrated that over expression of NHE-RF1 in CF airway cells induces a redistribution of ΔF508-CFTR from the cytoplasm to the apical membrane and rescues CFTR-dependent chloride secretion [155], effect reversibly observed in NHE-RF1-knockdown cells with increased degradation of the mutated protein [156]. In addition to direct interaction with CFTR, wild type over expression of

NHE-RF1 enhances the interaction between NHE-RF1, CFTR and ezrin, and ezrin with actin [157], these latter two also up regulated in CF's NEC. Ezrin is known to act as a protein kinase A (PKA)-anchoring protein and to associate with actin cytoskeleton [158] by connecting continuous filamentous (F)-actin to plasma membrane proteins aiding NHE-RF1 in stabilizing and maintain CFTR's function and activity [153,159,160]. NHE-RF1 is required for the maintenance of active ERM proteins at the cortical brush border membranes of polarized epithelia forming complexes with ERM and F-actin [161]. Interestingly, our results corroborate this cytoskeleton organization by differential expression of its constituents.

Besides ezrin and actin, several other proteins involved in cytoskeleton organization and maintenance have higher abundances in CF than in the other groups considered: TPM1, ACTN1, ACTN4, CAPZA1, CAPZB, PFN2 and EEF1G. CAPZA1, CAPZB, PFN1 and EEF1G also have higher abundances in patients with severe lung disease than mild ones (similar levels to healthy non-CF individuals) while PFN2 present higher levels in CF regardless lung function, supporting the enhanced tissue injury hypothesis. Conversely, levels of CFL1, VCL, FLNA and FLNB were similar in healthy and patients with mild lung disease and lower in a severe condition. These specimens are actin-interacting proteins controlling actin polymerization/despolymerization and its branching and attachment to plasma membrane. Filamins are also responsible for assembling actin to membrane glycoproteins, such as CFTR. Proper polarization status of the epithelial cells is highly dependent on adequate architectural arrangements of the cytoskeleton crucial for integrity and function of NE [162] and more specifically for cAMP-dependent CFTR activation [163]. Altered levels of proteins responsible for this assembly can serve as basis for abnormal remodeling in the CF airways.

Three other proteins involved in epithelial differentiation were found differentially expressed in this work: DMBT1, LGALS3 and LGALS3BP. DMBT1 is more abundant in CF carriers and patients than in healthy with similar pattern when respiratory function is accessed (more abundant in mild and severe patients than in non-CF). This protein may be considered as a candidate tumour suppressor gene for brain, lung, esophageal, gastric and colorectal cancers and play roles in mucosal defence system, cellular immune defence against bacterial pathogens by binding to a range of poly-sulfated and poly-phosphorylated ligands which may explain its broad bacterial-binding specificity. It associates with the actin cytoskeleton and, together with LGALS3

and LGALS3BP, is required for terminal differentiation of columnar epithelial cells during early embryogenesis. Galectin-3 presents higher abundance in CF-carriers and patients and is consistently up regulated from healthy to mild CF with the highest levels found in patients with severe lung disease, in accordance with the function performed by the protein.

Here, we were able to identify protein LPLUNC1 less abundant in CF patients with mild and severe lung disease in respective to controls. PLUNC's family are abundant proteins in the conducting airways estimated of making up about 1% of the total protein in human nasal lavage fluid (NLF) [164] and as much as 10% of the total protein in secretions from air-liquid interface cultures of human tracheobronchial epithelia [165] with proposed innate and host immune function. While the second condition seems to be verified, work to demonstrate PLUNC's activity is potent, broad-spectrum antimicrobial activity has failed. Previously results from Roxo-Rosa *et al.* (2006) reported PLUNC as not significantly up regulated in CF's NEC in a 2-DE experiment. Our results obtained by 2D-LC-MS/MS however show a different tendency. Although higher levels of PLUNC have been highlighted in lung-related conditions [166], novel roles for this family have been proposed. Secreted PLUNC inhibits sodium transport by ENaC in the air-liquid interface cultures of human airway epithelia, suggesting it to take part in ASL volume sensing and regulation [167]. Recent work from Gakhar *et al.* (2010) proved that physiological concentrations of PLUNC reduce surface tension in the airways to levels that enable immersion of inhaled bacteria, such as *P.aeruginosa*, and other particles into the ASL layer. This effect could maintain infecting bacteria in a susceptible state so they are not able to form infecting biofilms and be eradicated by other components of the innate immune system instead of a direct antimicrobial protein [168]. Similar outcomes exist for anti-biofilm host defences on mucosal surfaces, including in the airways. Lactoferrin (LTF) seems to play an important role in preventing biofilm formation by *P.aeruginosa* at concentrations below those that killed or prevent bacterial growth [169]. Recall that LTF was also less abundant in the membrane-enriched fraction of CF patients. Several of the patients enrolled for this study have their airways chronically colonized by persistent pathogens such as *P.aeruginosa* or *S.aureus*. Lower levels of LPLUNC1 and LTF can be closely related to ineffective inhibition of bacteria proliferation and their impaired clearance with consequences for the CF patients' lung disease.

3.3.2.1 Oxidative stress and reactive oxygen species generation

Chronic bacterial infection and inflammation lead to a self-perpetuating cycle of airway obstruction and destruction where oxidative stress and oxidant/antioxidant imbalance play a predominant role in the pathophysiology of CF. Also, ROS and RNS are described to be highly increased in this disease [170] with elevated levels of lipid and protein oxidation [171-174] and reduction of reduced glutathione levels (GSH). CFTR itself has been postulated to participate and regulate the apical efflux of GSH being an extra factor in the antioxidant status of epithelial lining fluid and an inherent cause of oxidative stress in CF [175-177]. Examples of imbalanced oxidative stress are the differential expression patterns of enzymes involved in protection and detoxification of airways. PRDX6 is a glutathione-dependent peroxidase constituting the major antioxidant defence in the lungs [178,179]. Levels of this protein were found similar among healthy groups and CF patients with differences when severity was concerned: PRDX6 decreased from healthy to mild CF patients and suddenly increased to severe patients overcoming the basal levels observed in controls. By a 2DE-based study, we previously observed a PRDX6 corresponding spot less abundant in CF's NEC probably as a specific isoform decreased in CF [16]. In accordance with that, recently Trudel *et al.* (2009) showed concordant evidences with increased levels of both PRDX6 mRNA and protein in *Cftr*$^{-/-}$ mice [180] and suggest that the PRDX6-dependent antioxidant capacity could be associated with post-translational modifications as three forms were identified, one of them exclusively in *Cftr*$^{+/+}$ mice after paraquat challenge. It could be the case that Roxo-Rosa *et al.* have also identified a specific isoform of PRDX6 as decreased in CF patients while our results and Trudel's suggest a stronger enzyme antioxidant response activated by PRDX6.

Other family of differentially expressed enzymes identified here are members of glutathione peroxidase and transferase family. GPX1 is a selenoprotein that catalyses the inactivation of ROS and RNS by oxidation of two GSH molecules thus protecting other biomolecules [181]. This enzyme presents its lowest level in CF carriers but also decreased expression in patients compared with healthy. Severe patients present higher levels of the enzyme than mild ones although similar amounts were observed in controls suggesting that GPX1 might constitute a first-line defence against oxidative damage in earlier stages of the disease.

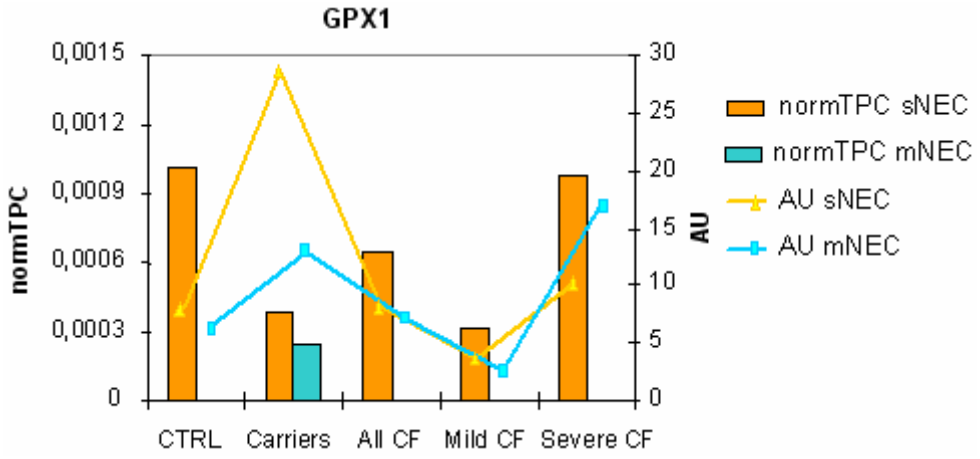


Figure V.5: Correlation between the normalized total peptide count (normTPC) and intensity band volume for GPX1 protein in a LC-MS/MS and Western Blot experiment respectively. (Orange bars: normTPC for GPX1 in the soluble fraction of NEC; Blue bars: normTPC for GPX1 in the membrane fraction of NEC; Orange line: arbibtraty units of the intensity of the band obtained from WB for GPX1 in the soluble fraction; Blue line: arbibtraty units of the intensity of the band obtained from WB for GPX1 in the membrane fraction)

GSTA1, GSTP1 and GSTK1 are members of the S-transferase superfamily that conjugate GSH with hydrophobic electrophiles to form conjugates for excretion [182]. GSTP1 evidenced similar levels in CF carriers and controls with decreased expression in patients, consistent with previous results obtained ate cellular [183] and tissue level [16], while GSTK1 showed only slightly increase from healthy to CF patients. When comparing lung disease severity, GSTA1 denoted an increase in mild patients with significant drop in severe ones. Although belong to the same family, GST polymorphisms have been associated with heat-shock response, electrophiles detoxification or immunomodulatory functions, among others [184]. Expression of members from this superfamily has also been investigated in different lung compartments of smokers and nonsmokers demonstrating tissue-dependent differential expression and specificity in xenobiotic metabolism [112]. A member of GST, GSTM1-0, has been proposed to significantly influence severity of CF lung disease [34], revealing important regulation of oxidant/antioxidant status in the phenotype. Other proteins involved in detoxification and balance of the oxidative status of cells are IDH1, PARK7 and AKR1B10, all with increased levels in CF, making undoubtedly the contribution of ROS and oxidative stress to CF's development and progression. However, that does not necessarily mean reduction on all antioxidant defences leading

to aggressive tissue damage as we and others postulate. The response of cells might lead in the sense of displaying a strong enzymatic antioxidant response differentially regulating the pathways.

3.3.2.2 Inflammation

Apolipoprotein APOA1, besides sustain the function of cardiovascular systems play also a critical protective role in preventing pulmonary inflammation, impaired vasodilatation and increased airway hyperresponsiveness [185].

Interestingly, only APOA1 showed consistent lower expression from healthy to CF carriers and from those to CF patients. APO2 presents lower levels in both CF and CF carriers in comparison with health individuals. We have recently demonstrated the same tendency of expression for these proteins in systemic circulation and attempted to link its reduced expression with CF and other respiratory diseases [186]. In human but not in mouse, APOA1 and APOA2 is activated by PPARα, a transcription factor involved in many biological process including down-regulation of pro-inflammatory genes and accordingly suppress the inflammatory response [187]. Reduction of both PPARα mRNA and protein amounts in lymphocytes of CF patients has been observed, which may render lymphocytes into cells that promote the inflammatory response and consequently lead to increased inflammation [188].

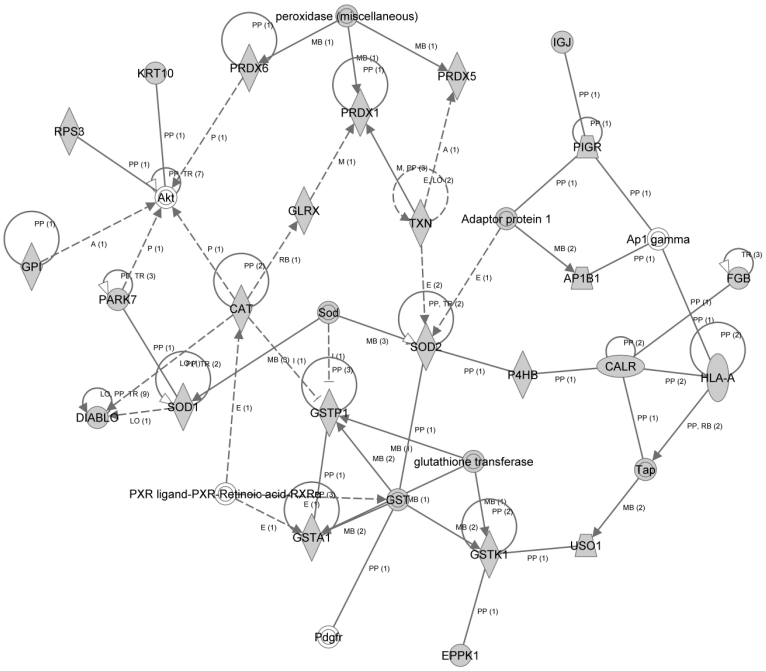
On the other hand, GOT2, a protein that we observed increased in CF carriers and higher in patients, was described to be activated by PPARα in mouse. Results in a murine small intestine model showed that GOT2 triggers activation of PPARα resulting in reduced cellular proliferation and increased differentiation of cells [189]. Either by stimulation or decreased activation, the study of similar behaviour at human airways remains to be elucidated but put strength on the role of these receptors in controlling phenomena intrinsically related with CF.

3.3.2.3 Functional Analysis

Although describing proteins at an individual level provide detailed and focused information on function and contribution to the disease, functional analysis of the differentially expressed proteins allows their integration into networks and pathways with biological significance (Tables V.5SDV.6SD and Figures V.3SD and V.4SD). Regulation of the proteins participating in such processes might be a different approach

to get a broad understanding of the mechanisms underlined in CF. Pathway analysis corroborate our results of the individual proteins scrutiny by emphasizing glutathione and xenobiotics metabolism, oxidative stress response, IL-8 signalling or PPARα activation as deregulated processes in CF. Even ubiquitous processes such as glycolysis and gluconeogeneis, the citrate cycle or pentose phosphate pathway are highlighted as imbalanced in CF and can be interpreted as deficit in generation of reducing equivalents and energy production resulting in impaired metabolic profile of these cells [190,191].

Differentially expressed proteins were allocated to representative networks of the processes ascribed before with especial emphasis in small molecule and drug metabolism, cellular assembly and organization, immune response or infection mechanisms (Figure V.6A). Interestingly, one of the networks retrieved was associated with respiratory disease and featured CFTR that, although not identified, was clearly associated with impaired function of the lungs (Figure V.6B).



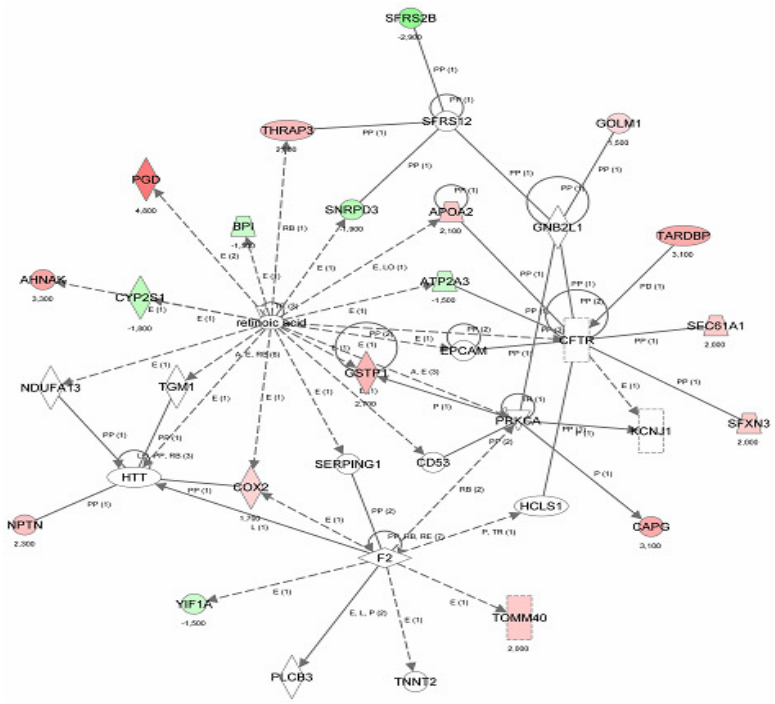


Figure V.6: Illustrative networks related to antioxidant status (**6A**) and respiratory function (**6B**), retrieved by IPA's knowledgebase.

SUMMARY

We have showed that nasal epithelial cells (NEC) can be used as a surrogate model of the lower airways and therefore provide deeper insights into the mechanisms underlying the pathogenesis of CF lung disease.

The combination of an improved methodology of cell fractionation and separation of proteins from NEC allowed the broad identification and characterization of differentially expressed proteins with significant implications in the progression of CF lung disease. From CFTR-interacting proteins to CF-modifier genes, several specimens described here have putative roles in the development and/or progression of CF lung disease. Integration of these proteins in pathways and networks with biological significance corroborate our findings at the protein level, retrieving known impaired functions already associated with CF, such as tissue remodeling, inflammation or antioxidant imbalance, but also providing glances at other new processes like mitochondrial dysfunction, impaired energy production and altered glucose's metabolism.

We believe that our findings constitute a solid base for more focused and dedicated studies on relevant proteins highlighted here that ultimately can constitute potential therapeutic and/or auxiliary diagnostic tools for CF.

REFERENCES

[1] Cystic Fibrosis Foundation. Cystic Fibrosis Foundation, Patient Registry 2005 Annual Report, Bethesda, MD. 2004.

[2] Welsh MJ, Smith AE. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. Cell 1993; 73:1251-4.

[3] Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in cystic fibrosis. Am J Respir Crit Care Med 2003; 168:918-51.

[4] Ramsey BW. Management of pulmonary disease in patients with cystic fibrosis. N Engl J Med 1996; 335:179-88.

[5] MacGregor G, Gray RD, Hilliard TN, Imrie M, Boyd AC, *et al.* Biomarkers for cystic fibrosis lung disease: application of SELDI-TOF mass spectrometry to BAL fluid. J Cyst Fibros 2008; 7:352-8.

[6] von Bredow C, Birrer P, Griese M. Surfactant protein A and other bronchoalveolar lavage fluid proteins are altered in cystic fibrosis. Eur Respir J 2001; 17:716-22.

[7] Gharib SA, Vaisar T, Aitken ML, Park DR, Heinecke JW, *et al.* Mapping the lung proteome in cystic fibrosis. J Proteome Res 2009; 8:3020-8.

[8] Sloane AJ, Lindner RA, Prasad SS, Sebastian LT, Pedersen SK, *et al.* Proteomic analysis of sputum from adults and children with cystic fibrosis and from control subjects. Am J Respir Crit Care Med 2005; 172:1416-26.

[9] Pedersen SK, Sloane AJ, Prasad SS, Sebastian LT, Lindner RA, *et al.* An immunoproteomic approach for identification of clinical biomarkers for monitoring disease: application to cystic fibrosis. Mol Cell Proteomics 2005; 4:1052-60.

[10] Gray RD, Imrie M, Boyd AC, Porteous D, Innes JA, *et al.* Sputum and serum calprotectin are useful biomarkers during CF exacerbation. J Cyst Fibros 2010; 9:193-8.

[11] Gray RD, Duncan A, Noble D, Imrie M, O'Reilly DS, *et al.* Sputum trace metals are biomarkers of inflammatory and suppurative lung disease. Chest 2010; 137:635-41.

[12] Inglis SK, Wilson SM. Cystic fibrosis and airway submucosal glands. Pediatr Pulmonol 2005; 40:279-84.

[13] Donaldson SH, Bennett WD, Zeman KL, Knowles MR, Tarran R, *et al.* Mucus clearance and lung function in cystic fibrosis with hypertonic saline. N Engl J Med 2006; 354:241-50.

[14] Blouquit S, Regnier A, Dannhoffer L, Fermanian C, Naline E, *et al.* Ion and fluid transport properties of small airways in cystic fibrosis. Am J Respir Crit Care Med 2006; 174:299-305.

[15] Matsui H, Grubb BR, Tarran R, Randell SH, Gatzy JT, *et al.* Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease. Cell 1998; 95:1005-15.

[16] Roxo-Rosa M, da Costa G, Luider TM, Scholte BJ, Coelho AV, *et al.* Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006; 6:2314-25.

[17] Wright JM, Merlo CA, Reynolds JB, Zeitlin PL, Garcia JG, *et al.* Respiratory epithelial gene expression in patients with mild and severe cystic fibrosis lung disease. Am J Respir Cell Mol Biol 2006; 35:327-36.

[18] Gomes-Alves P, Imrie M, Gray RD, Nogueira P, Ciordia S*, et al.* SELDI-TOF biomarker signatures for cystic fibrosis, asthma and chronic obstructive pulmonary disease. Clin Biochem 2010; 43:168-77.
[19] Pier GB. Airway epithelial (nasal) cell monolayers used to study Pseudomonas aeruginosa invasion are hyperpolarized and not representative of the human airway epithelium. Infect Immun 2006; 74:7043; author reply 7043-4.
[20] Gaga M, Vignola AM, Chanez P. Upper and lower airways: similarities and differences. Eur Respir Mon 2001; 18:1-15.
[21] Slavin RG. The upper and lower airways: the epidemiological and pathophysiological connection. Allergy Asthma Proc 2008; 29:553-6.
[22] McDougall CM, Blaylock MG, Douglas JG, Brooker RJ, Helms PJ*, et al.* Nasal epithelial cells as surrogates for bronchial epithelial cells in airway inflammation studies. Am J Respir Cell Mol Biol 2008; 39:560-8.
[23] Bonestroo HJ, de Winter-de Groot KM, van der Ent CK, Arets HG. Upper and lower airway cultures in children with cystic fibrosis: Do not neglect the upper airways. J Cyst Fibros 2010; 9:130-134.
[24] Hurst JR, Wilkinson TM, Perera WR, Donaldson GC, Wedzicha JA. Relationships among bacteria, upper airway, lower airway, and systemic inflammation in COPD. Chest 2005; 127:1219-26.
[25] Zulianello L, Canard C, Kohler T, Caille D, Lacroix JS*, et al.* Rhamnolipids are virulence factors that promote early infiltration of primary human airway epithelia by Pseudomonas aeruginosa. Infect Immun 2006; 74:3134-47.
[26] Togias A. Systemic cross-talk between the lung and the nose. Am J Respir Crit Care Med 2001; 164:726-7.
[27] Beck S, Penque D, Garcia S, Gomes A, Farinha C*, et al.* Cystic fibrosis patients with the 3272-26A-->G mutation have mild disease, leaky alternative mRNA splicing, and CFTR protein at the cell membrane. Hum Mutat 1999; 14:133-44.
[28] Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R*, et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989; 245:1066-73.
[29] Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK*, et al.* Identification of the cystic fibrosis gene: genetic analysis. Science 1989; 245:1073-80.
[30] Kerem E, Corey M, Kerem BS, Rommens J, Markiewicz D*, et al.* The relation between genotype and phenotype in cystic fibrosis--analysis of the most common mutation (delta F508). N Engl J Med 1990; 323:1517-22.
[31] Kerem B, Kerem E. The molecular basis for disease variability in cystic fibrosis. Eur J Hum Genet 1996; 4:65-73.
[32] Mickle JE, Cutting GR. Genotype-phenotype relationships in cystic fibrosis. Med Clin North Am 2000; 84:597-607.
[33] Hull J, Thomson AH. Contribution of genetic factors other than CFTR to disease severity in cystic fibrosis. Thorax 1998; 53:1018-21.
[34] Merlo CA, Boyle MP. Modifier genes in cystic fibrosis lung disease. J Lab Clin Med 2003; 141:237-41.
[35] Griesenbach U, Alton EWFW, Davies J. CF Modifier Genes. in: A. Bush, E.W.F.W. Alton, J.C. Davies, U. Griesenbach, and A. Jaffe, (Eds.), Cystic fibrosis in the 21st century, Karger, Basel, 2006, pp. 77-83.
[36] Penque D, Mendes F, Beck S, Farinha C, Pacheco P*, et al.* Cystic fibrosis F508del patients have apically localized CFTR in a reduced number of airway cells. Lab Invest 2000; 80:857-68.

[37] Johnson C, Butler SM, Konstan MW, Morgan W, Wohl ME. Factors influencing outcomes in cystic fibrosis: a center-based analysis. Chest 2003; 123:20-7.
[38] Cystic Fibrosis Foundation, Cystic Fibrosis Mutation Database, Bethesda, 1989.
[39] Pacheco P, Detection rate of CF causing CFTR mutations in Portugal. in: WHO, (Ed.), The Molecular Genetic Epidemiology of Cystic Fibrosis, Human Genetics Programme, Chronic Diseases and Health Promotion, 2004.
[40] Carvalho PC, Hewel J, Barbosa VC, Yates JR, 3rd. Identifying differences in protein expression levels by spectral counting and feature selection. Genet Mol Res 2008; 7:342-56.
[41] Rappsilber J, Ryder U, Lamond AI, Mann M. Large-scale proteomic analysis of the human spliceosome. Genome Res 2002; 12:1231-45.
[42] Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T*, et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 2005; 4:1265-72.
[43] Zabner J, Scheetz TE, Almabrazi HG, Casavant TL, Huang J*, et al.* CFTR DeltaF508 mutation has minimal effect on the gene expression profile of differentiated human airway epithelia. Am J Physiol Lung Cell Mol Physiol 2005; 289:L545-53.
[44] Carvalho-Oliveira IM, Charro N, Aarbiou J, Buijs-Offerman RM, Wilke M*, et al.* Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009; 8:3606-16.
[45] Xu Y, Clark JC, Aronow BJ, Dey CR, Liu C*, et al.* Transcriptional adaptation to cystic fibrosis transmembrane conductance regulator deficiency. J Biol Chem 2003; 278:7674-82.
[46] Miller I, Eberini I, Gianazza E. Proteomics of lung physiopathology. Proteomics 2008; 8:5053-73.
[47] Hosking CR, Ulloa F, Hogan C, Ferber EC, Figueroa A*, et al.* The transcriptional repressor Glis2 is a novel binding partner for p120 catenin. Mol Biol Cell 2007; 18:1918-27.
[48] Daniel JM, Reynolds AB. The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. Mol Cell Biol 1999; 19:3614-23.
[49] Chai JY, Modak C, Mouazzen W, Narvaez R, Pham J. Epithelial or mesenchymal: Where to draw the line? Biosci Trends 2010; 4:130-42.
[50] Castillo SD, Angulo B, Suarez-Gauthier A, Melchor L, Medina PP*, et al.* Gene amplification of the transcription factor DP1 and CTNND1 in human lung cancer. J Pathol 2010.
[51] Kondapalli J, Flozak AS, Albuquerque ML. Laminar shear stress differentially modulates gene expression of p120 catenin, Kaiso transcription factor, and vascular endothelial cadherin in human coronary artery endothelial cells. J Biol Chem 2004; 279:11417-24.
[52] Zhang Y, Wolf-Yadlin A, Ross PL, Pappin DJ, Rush J*, et al.* Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. Mol Cell Proteomics 2005; 4:1240-50.
[53] Rikova K, Guo A, Zeng Q, Possemato A, Yu J*, et al.* Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell 2007; 131:1190-203.

[54] Fliegauf M, Benzing T, Omran H. When cilia go bad: cilia defects and ciliopathies. Nat Rev Mol Cell Biol 2007; 8:880-93.
[55] Poon TC, Pang RT, Chan KC, Lee NL, Chiu RW*, et al.* Proteomic profiling in SARS: diagnostic and prognostic applications. Hong Kong Med J 2009; 15 Suppl 8:15-8.
[56] Chang JT, Chang GC, Ko JL, Liao HY, Liu HJ*, et al.* Induction of tubulin by docetaxel is associated with p53 status in human non small cell lung cancer cell lines. Int J Cancer 2006; 118:317-25.
[57] Chu JJ, Chiang CD, Rao CS, Chang WM, Lai YK. Establishment and characterization of a paclitaxel-resistant human non-small cell lung cancer cell line. Anticancer Res 2000; 20:2449-56.
[58] Moll R, Divo M, Langbein L. The human keratins: biology and pathology. Histochem Cell Biol 2008; 129:705-33.
[59] Upasani OS, Vaidya MM, Bhisey AN. Database on monoclonal antibodies to cytokeratins. Oral Oncol 2004; 40:236-56.
[60] Ku NO, Darling JM, Krams SM, Esquivel CO, Keeffe EB*, et al.* Keratin 8 and 18 mutations are risk factors for developing liver disease of multiple etiologies. Proc Natl Acad Sci U S A 2003; 100:6063-8.
[61] Zielenski J. Genotype and phenotype in cystic fibrosis. Respiration 2000; 67:117-33.
[62] Zielenski J, Tsui LC. Cystic fibrosis: genotypic and phenotypic variations. Annu Rev Genet 1995; 29:777-807.
[63] Schlage WK, Bulles H, Friedrichs D, Kuhn M, Teredesai A. Cytokeratin expression patterns in the rat respiratory tract as markers of epithelial differentiation in inhalation toxicology. I. Determination of normal cytokeratin expression patterns in nose, larynx, trachea, and lung. Toxicol Pathol 1998; 26:324-43.
[64] Davezac N, Tondelier D, Lipecka J, Fanen P, Demaugre F*, et al.* Global proteomic approach unmasks involvement of keratins 8 and 18 in the delivery of cystic fibrosis transmembrane conductance regulator (CFTR)/deltaF508-CFTR to the plasma membrane. Proteomics 2004; 4:3833-44.
[65] Lipecka J, Norez C, Bensalem N, Baudouin-Legros M, Planelles G*, et al.* Rescue of DeltaF508-CFTR (cystic fibrosis transmembrane conductance regulator) by curcumin: involvement of the keratin 18 network. J Pharmacol Exp Ther 2006; 317:500-5.
[66] Lanas A, Ferrandez A. Inappropriate prevention of NSAID-induced gastrointestinal events among long-term users in the elderly. Drugs Aging 2007; 24:121-31.
[67] Oliver MN, Fabry B, Marinkovic A, Mijailovich SM, Butler JP*, et al.* Airway hyperresponsiveness, remodeling, and smooth muscle mass: right answer, wrong reason? Am J Respir Cell Mol Biol 2007; 37:264-72.
[68] Tiddens HA, Donaldson SH, Rosenfeld M, Pare PD. Cystic fibrosis lung disease starts in the small airways: can we treat it more effectively? Pediatr Pulmonol 2010; 45:107-17.
[69] Orian-Rousseau V. CD44, a therapeutic target for metastasising tumours. Eur J Cancer 2010; 46:1271-7.
[70] Johnson P, Ruffell B. CD44 and its role in inflammation and inflammatory diseases. Inflamm Allergy Drug Targets 2009; 8:208-20.
[71] Nusgens BV. [Hyaluronic acid and extracellular matrix: a primitive molecule?]. Ann Dermatol Venereol 2010; 137 Suppl 1:S3-8.

[72] Meezaman D, Charles P, Daskal E, Polymeropoulos MH, Martin BM*, et al.* Cloning and analysis of cDNA encoding a major airway glycoprotein, human tracheobronchial mucin (MUC5). J Biol Chem 1994; 269:12932-9.
[73] Van de Bovenkamp JH, Mahdavi J, Korteland-Van Male AM, Buller HA, Einerhand AW*, et al.* The MUC5AC glycoprotein is the primary receptor for Helicobacter pylori in the human stomach. Helicobacter 2003; 8:521-32.
[74] Silverman HS, Parry S, Sutton-Smith M, Burdick MD, McDermott K*, et al.* In vivo glycosylation of mucin tandem repeats. Glycobiology 2001; 11:459-71.
[75] Perez-Vilar J, Randell SH, Boucher RC. C-Mannosylation of MUC5AC and MUC5B Cys subdomains. Glycobiology 2004; 14:325-37.
[76] Verkman AS, Song Y, Thiagarajah JR. Role of airway surface liquid and submucosal glands in cystic fibrosis lung disease. Am J Physiol Cell Physiol 2003; 284:C2-15.
[77] Verkman AS. Physiological importance of aquaporins: lessons from knockout mice. Curr Opin Nephrol Hypertens 2000; 9:517-22.
[78] Matsuzaki T, Hata H, Ozawa H, Takata K. Immunohistochemical localization of the aquaporins AQP1, AQP3, AQP4, and AQP5 in the mouse respiratory system. Acta Histochem Cytochem 2009; 42:159-69.
[79] Shen Y, Wang Y, Chen Z, Wang D, Wang X*, et al.* Role of Aquaporin 5 in antigen-induced airway inflammation and mucous hyperproduction in mice. J Cell Mol Med 2010.
[80] Chen Z, Wang X, Gao L, Bai L, Zhu R*, et al.* Regulation of MUC5AC mucin secretion by depletion of AQP5 in SPC-A1 cells. Biochem Biophys Res Commun 2006; 342:775-81.
[81] Vroling AB, Fokkens WJ, van Drunen CM. How epithelial cells detect danger: aiding the immune response. Allergy 2008; 63:1110-23.
[82] Baltimore RS, Christie CD, Smith GJ. Immunohistopathologic localization of Pseudomonas aeruginosa in lungs from patients with cystic fibrosis. Implications for the pathogenesis of progressive lung deterioration. Am Rev Respir Dis 1989; 140:1650-61.
[83] Thomson E, Brennan S, Senthilmohan R, Gangell CL, Chapman AL*, et al.* Identifying Peroxidases and Their Oxidants in the Early Pathology of Cystic Fibrosis. Free Radic Biol Med 2010.
[84] Wilde CG, Seilhamer JJ, McGrogan M, Ashton N, Snable JL*, et al.* Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-binding protein. LPS binding properties and effects on LPS-mediated cell activation. J Biol Chem 1994; 269:17411-6.
[85] Wasiluk KR, Skubitz KM, Gray BH. Comparison of granule proteins from human polymorphonuclear leukocytes which are bactericidal toward Pseudomonas aeruginosa. Infect Immun 1991; 59:4193-200.
[86] Witko-Sarsat V, Delacourt C, Rabier D, Bardet J, Nguyen AT*, et al.* Neutrophil-derived long-lived oxidants in cystic fibrosis sputum. Am J Respir Crit Care Med 1995; 152:1910-6.
[87] Vandivier RW, Fadok VA, Hoffmann PR, Bratton DL, Penvari C*, et al.* Elastase-mediated phosphatidylserine receptor cleavage impairs apoptotic cell clearance in cystic fibrosis and bronchiectasis. J Clin Invest 2002; 109:661-70.
[88] Witko-Sarsat V, Halbwachs-Mecarelli L, Schuster A, Nusbaum P, Ueki I*, et al.* Proteinase 3, a potent secretagogue in airways, is present in cystic fibrosis sputum. Am J Respir Cell Mol Biol 1999; 20:729-36.

[89] Ordonez CL, Henig NR, Mayer-Hamblett N, Accurso FJ, Burns JL*, et al.* Inflammatory and microbiologic markers in induced sputum after intravenous antibiotics in cystic fibrosis. Am J Respir Crit Care Med 2003; 168:1471-5.
[90] Sagel SD, Sontag MK, Wagener JS, Kapsner RK, Osberg I*, et al.* Induced sputum inflammatory measures correlate with lung function in children with cystic fibrosis. J Pediatr 2002; 141:811-7.
[91] Husby G. Is there a pathogenic link between gamma heavy chain disease and chronic arthritis? Curr Opin Rheumatol 2000; 12:65-70.
[92] von Pawel-Rammingen U, Johansson BP, Bjorck L. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. Embo J 2002; 21:1607-15.
[93] Pyles RB, Jezek GE, Eaves-Pyles TD. Toll-like receptor 3 agonist protection against experimental Francisella tularensis respiratory tract infection. Infect Immun 2010; 78:1700-10.
[94] Cristofaro P, Opal SM. Role of Toll-like receptors in infection and immunity: clinical implications. Drugs 2006; 66:15-29.
[95] Homma T, Matsukura S, Hirose T, Ohnishi T, Kimura T*, et al.* Cooperative activation of CCL5 expression by TLR3 and tumor necrosis factor-alpha or interferon-gamma through nuclear factor-kappaB or STAT-1 in airway epithelial cells. Int Arch Allergy Immunol 2010; 152 Suppl 1:9-17.
[96] Teng F, Slavik V, Duffy KE, San Mateo L, Goldschmidt R. Toll-like receptor 3 is involved in airway epithelial cell response to nontypeable Haemophilus influenzae. Cell Immunol 2010; 260:98-104.
[97] Rudd BD, Smit JJ, Flavell RA, Alexopoulou L, Schaller MA*, et al.* Deletion of TLR3 alters the pulmonary immune environment and mucus production during respiratory syncytial virus infection. J Immunol 2006; 176:1937-42.
[98] Gleich GJ. Mechanisms of eosinophil-associated inflammation. J Allergy Clin Immunol 2000; 105:651-63.
[99] Wu W, Chen Y, d'Avignon A, Hazen SL. 3-Bromotyrosine and 3,5-dibromotyrosine are major products of protein oxidation by eosinophil peroxidase: potential markers for eosinophil-dependent tissue injury in vivo. Biochemistry 1999; 38:3538-48.
[100] Denzler KL, Borchers MT, Crosby JR, Cieslewicz G, Hines EM*, et al.* Extensive eosinophil degranulation and peroxidase-mediated oxidation of airway proteins do not occur in a mouse ovalbumin-challenge model of pulmonary inflammation. J Immunol 2001; 167:1672-82.
[101] MacPherson JC, Comhair SA, Erzurum SC, Klein DF, Lipscomb MF*, et al.* Eosinophils are a major source of nitric oxide-derived oxidants in severe asthma: characterization of pathways available to eosinophils for generating reactive nitrogen species. J Immunol 2001; 166:5763-72.
[102] Jong EC, Klebanoff SJ. Eosinophil-mediated mammalian tumor cell cytotoxicity: role of the peroxidase system. J Immunol 1980; 124:1949-53.
[103] Jong EC, Chi EY, Klebanoff SJ. Human neutrophil-mediated killing of schistosomula of Schistosoma mansoni: augmentation by schistosomal binding of eosinophil peroxidase. Am J Trop Med Hyg 1984; 33:104-15.
[104] Klebanoff SJ, Coombs RW. Virucidal effect of stimulated eosinophils on human immunodeficiency virus type 1. AIDS Res Hum Retroviruses 1996; 12:25-9.
[105] DeChatelet LR, Shirley PS, McPhail LC, Huntley CC, Muss HB*, et al.* Oxidative metabolism of the human eosinophil. Blood 1977; 50:525-35.

[106] Agosti JM, Altman LC, Ayars GH, Loegering DA, Gleich GJ*, et al.* The injurious effect of eosinophil peroxidase, hydrogen peroxide, and halides on pneumocytes in vitro. J Allergy Clin Immunol 1987; 79:496-504.
[107] McElhinney B, Poynter ME, Shrivastava P, Hazen SL, Janssen-Heininger YM. Eosinophil peroxidase catalyzes JNK-mediated membrane blebbing in a Rho kinase-dependent manner. J Leukoc Biol 2003; 74:897-907.
[108] Crul T, Testelmans D, Spruit MA, Troosters T, Gosselink R*, et al.* Gene expression profiling in vastus lateralis muscle during an acute exacerbation of COPD. Cell Physiol Biochem 2010; 25:491-500.
[109] Yusenko MV, Ruppert T, Kovacs G. Analysis of differentially expressed mitochondrial proteins in chromophobe renal cell carcinomas and renal oncocytomas by 2-D gel electrophoresis. Int J Biol Sci 2010; 6:213-24.
[110] Seth R, Keeley J, Abu-Ali G, Crook S, Jackson D*, et al.* The putative tumour modifier gene ATP5A1 is not mutated in human colorectal cancer cell lines but expression levels correlate with TP53 mutations and chromosomal instability. J Clin Pathol 2009; 62:598-603.
[111] Saarikoski ST, Rivera SP, Hankinson O, Husgafvel-Pursiainen K. CYP2S1: a short review. Toxicol Appl Pharmacol 2005; 207:62-9.
[112] Thum T, Erpenbeck VJ, Moeller J, Hohlfeld JM, Krug N*, et al.* Expression of xenobiotic metabolizing enzymes in different lung compartments of smokers and nonsmokers. Environ Health Perspect 2006; 114:1655-61.
[113] Nishida CR, Lee M, Ortiz De Montellano PR. Efficient Hypoxic Activation of the Anticancer Agent AQ4N by CYP2S1 and CYP2W1. Mol Pharmacol 2010.
[114] Pavez Lorie E, Li H, Vahlquist A, Torma H. The involvement of cytochrome p450 (CYP) 26 in the retinoic acid metabolism of human epidermal keratinocytes. Biochim Biophys Acta 2009; 1791:220-8.
[115] Mark M, Ghyselinck NB, Chambon P. Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. Annu Rev Pharmacol Toxicol 2006; 46:451-80.
[116] Guilbault C, De Sanctis JB, Wojewodka G, Saeed Z, Lachance C*, et al.* Fenretinide corrects newly found ceramide deficiency in cystic fibrosis. Am J Respir Cell Mol Biol 2008; 38:47-56.
[117] Saeed Z, Guilbault C, De Sanctis JB, Henri J, Marion D*, et al.* Fenretinide prevents the development of osteoporosis in Cftr-KO mice. J Cyst Fibros 2008; 7:222-30.
[118] Doherty GJ, McMahon HT. Mechanisms of endocytosis. Annu Rev Biochem 2009; 78:857-902.
[119] Sottile J, Chandler J. Fibronectin matrix turnover occurs through a caveolin-1-dependent process. Mol Biol Cell 2005; 16:757-68.
[120] Fustik S, Jacovska T, Spirevska L, Koceva S. Protein-energy malnutrition as the first manifestation of cystic fibrosis in infancy. Pediatr Int 2009; 51:678-83.
[121] Com G, Clancy JP. Adenosine receptors, cystic fibrosis, and airway hydration. Handb Exp Pharmacol 2009; 363-81.
[122] Zhang H, Wang G, Ding Y, Wang Z, Barraclough R*, et al.* The crystal structure at 2A resolution of the Ca2+ -binding protein S100P. J Mol Biol 2003; 325:785-94.
[123] Arumugam T, Logsdon CD. S100P: a novel therapeutic target for cancer. Amino Acids 2011.

[124] Diederichs S, Bulk E, Steffen B, Ji P, Tickenbrock L*, et al.* S100 family members and trypsinogens are predictors of distant metastasis and survival in early-stage non-small cell lung cancer. Cancer Res 2004; 64:5564-9.
[125] Kim B, Lee HJ, Choi HY, Shin Y, Nam S*, et al.* Clinical validity of the lung cancer biomarkers identified by bioinformatics analysis of public expression data. Cancer Res 2007; 67:7431-8.
[126] Bartling B, Rehbein G, Schmitt WD, Hofmann HS, Silber RE*, et al.* S100A2-S100P expression profile and diagnosis of non-small cell lung carcinoma: impairment by advanced tumour stages and neoadjuvant chemotherapy. Eur J Cancer 2007; 43:1935-43.
[127] Bulk E, Hascher A, Liersch R, Mesters RM, Diederichs S*, et al.* Adjuvant therapy with small hairpin RNA interference prevents non-small cell lung cancer metastasis development in mice. Cancer Res 2008; 68:1896-904.
[128] Arumugam T, Simeone DM, Schmidt AM, Logsdon CD. S100P stimulates cell proliferation and survival via receptor for activated glycation end products (RAGE). J Biol Chem 2004; 279:5059-65.
[129] Welsh MJ, Smith AE. Cystic fibrosis. Sci Am 1995; 273:52-9.
[130] Srivastava M, Eidelman O, Jozwik C, Paweletz C, Huang W*, et al.* Serum proteomic signature for cystic fibrosis using an antibody microarray platform. Mol Genet Metab 2006; 87:303-10.
[131] Lee IT, Lin CC, Wu YC, Yang CM. TNF-alpha induces matrix metalloproteinase-9 expression in A549 cells: role of TNFR1/TRAF2/PKCalpha-dependent signaling pathways. J Cell Physiol 2010; 224:454-64.
[132] Pinto LA, Depner M, Klopp N, Illig T, Vogelberg C*, et al.* MMP-9 gene variants increase the risk for non-atopic asthma in children. Respir Res 2010; 11:23.
[133] Sampsonas F, Kaparianos A, Lykouras D, Karkoulias K, Spiropoulos K. DNA sequence variations of metalloproteinases: their role in asthma and COPD. Postgrad Med J 2007; 83:244-50.
[134] Wallace AM, Sandford AJ. Genetic polymorphisms of matrix metalloproteinases: functional importance in the development of chronic obstructive pulmonary disease? Am J Pharmacogenomics 2002; 2:167-75.
[135] Zheng S, Chang Y, Hodges KB, Sun Y, Ma X*, et al.* Expression of KISS1 and MMP-9 in non-small cell lung cancer and their relations to metastasis and survival. Anticancer Res 2010; 30:713-8.
[136] Tabary O, Zahm JM, Hinnrasky J, Couetil JP, Cornillet P*, et al.* Selective up-regulation of chemokine IL-8 expression in cystic fibrosis bronchial gland cells in vivo and in vitro. Am J Pathol 1998; 153:921-30.
[137] Schmitt-Grohe S, Naujoks C, Bargon J, Wagner TO, Schubert R*, et al.* Interleukin-8 in whole blood and clinical status in cystic fibrosis. Cytokine 2005; 29:18-23.
[138] Pitrez PM, Brennan S, Turner S, Sly PD. Nasal wash as an alternative to bronchoalveolar lavage in detecting early pulmonary inflammation in children with cystic fibrosis. Respirology 2005; 10:177-82.
[139] Wallace AM, Sandford AJ, English JC, Burkett KM, Li H*, et al.* Matrix metalloproteinase expression by human alveolar macrophages in relation to emphysema. Copd 2008; 5:13-23.
[140] Chapman HA, Riese RJ, Shi GP. Emerging roles for cysteine proteases in human biology. Annu Rev Physiol 1997; 59:63-88.

[141] Kos J, Sekirnik A, Kopitar G, Cimerman N, Kayser K, *et al.* Cathepsin S in tumours, regional lymph nodes and sera of patients with lung cancer: relation to prognosis. Br J Cancer 2001; 85:1193-200.
[142] Cimerman N, Brguljan PM, Krasovec M, Suskovic S, Kos J. Circadian and concentration profile of cathepsin S in sera from healthy subjects and asthmatic patients. Pflugers Arch 2001; 442:R204-6.
[143] Sediva A, Bartunkova J, Kolarova I, Hrusak O, Vavrova V, *et al.* Antineutrophil cytoplasmic autoantibodies (ANCA) in children with cystic fibrosis. J Autoimmun 1998; 11:185-90.
[144] Sediva A, Lischke R, Simonek J, Tkaczyk J, Vavrova V, *et al.* Lung transplantation for cystic fibrosis: immune system and autoimmunity. Med Sci Monit 2001; 7:1219-23.
[145] Jensen P, Johansen HK, Carmi P, Hoiby N, Cohen IR. Autoantibodies to pancreatic hsp60 precede the development of glucose intolerance in patients with cystic fibrosis. J Autoimmun 2001; 17:165-72.
[146] Lachenal F, Nkana K, Nove-Josserand R, Fabien N, Durieu I. Prevalence and clinical significance of auto-antibodies in adults with cystic fibrosis. Eur Respir J 2009; 34:1079-85.
[147] Uhlen M, Bjorling E, Agaton C, Szigyarto CA, Amini B, *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 2005; 4:1920-32.
[148] Stockley RA, Bayley D, Hill SL, Hill AT, Crooks S, *et al.* Assessment of airway neutrophils by sputum colour: correlation with airways inflammation. Thorax 2001; 56:366-72.
[149] Laszlo A, Sohar I, Karacsonyi S, Petri A, Trojan I. Activities of serum cathepsin (B, H and L) and metalloproteinase (MMP7-ase) in patients with gastrointestinal and bronchial malignant tumours. Acta Med Hung 1990; 47:107-9.
[150] Sevenich L, Schurigt U, Sachse K, Gajda M, Werner F, *et al.* Synergistic antitumor effects of combined cathepsin B and cathepsin Z deficiencies on breast cancer progression and metastasis in mice. Proc Natl Acad Sci U S A 2010; 107:2497-502.
[151] Hall RA, Ostedgaard LS, Premont RT, Blitzer JT, Rahman N, *et al.* A C-terminal motif found in the beta2-adrenergic receptor, P2Y1 receptor and cystic fibrosis transmembrane conductance regulator determines binding to the Na+/H+ exchanger regulatory factor family of PDZ proteins. Proc Natl Acad Sci U S A 1998; 95:8496-501.
[152] Moyer BD, Duhaime M, Shaw C, Denton J, Reynolds D, *et al.* The PDZ-interacting domain of cystic fibrosis transmembrane conductance regulator is required for functional expression in the apical plasma membrane. J Biol Chem 2000; 275:27069-74.
[153] Short DB, Trotter KW, Reczek D, Kreda SM, Bretscher A, *et al.* An apical PDZ protein anchors the cystic fibrosis transmembrane conductance regulator to the cytoskeleton. J Biol Chem 1998; 273:19797-801.
[154] Seidler U, Singh A, Chen M, Cinar A, Bachmann O, *et al.* Knockout mouse models for intestinal electrolyte transporters and regulatory PDZ adaptors: new insights into cystic fibrosis, secretory diarrhoea and fructose-induced hypertension. Exp Physiol 2009; 94:175-9.
[155] Fanelli T, Cardone RA, Favia M, Guerra L, Zaccolo M, *et al.* Beta-oestradiol rescues DeltaF508CFTR functional expression in human cystic fibrosis airway

CFBE41o- cells through the up-regulation of NHERF1. Biol Cell 2008; 100:399-412.
[156] Kwon SH, Pollard H, Guggino WB. Knockdown of NHERF1 enhances degradation of temperature rescued DeltaF508 CFTR from the cell surface of human airway cells. Cell Physiol Biochem 2007; 20:763-72.
[157] Favia M, Guerra L, Fanelli T, Cardone RA, Monterisi S*, et al.* Na+/H+ exchanger regulatory factor 1 overexpression-dependent increase of cytoskeleton organization is fundamental in the rescue of F508del cystic fibrosis transmembrane conductance regulator in human airway CFBE41o- cells. Mol Biol Cell 2010; 21:73-86.
[158] Dransfield DT, Bradford AJ, Smith J, Martin M, Roy C*, et al.* Ezrin is a cyclic AMP-dependent protein kinase anchoring protein. Embo J 1997; 16:35-43.
[159] Swiatecka-Urban A, Duhaime M, Coutermarsh B, Karlson KH, Collawn J*, et al.* PDZ domain interaction controls the endocytic recycling of the cystic fibrosis transmembrane conductance regulator. J Biol Chem 2002; 277:40099-105.
[160] Turunen O, Wahlstrom T, Vaheri A. Ezrin has a COOH-terminal actin-binding site that is conserved in the ezrin protein family. J Cell Biol 1994; 126:1445-53.
[161] Morales FC, Takahashi Y, Kreimann EL, Georgescu MM. Ezrin-radixin-moesin (ERM)-binding phosphoprotein 50 organizes ERM proteins at the apical membrane of polarized epithelia. Proc Natl Acad Sci U S A 2004; 101:17705-10.
[162] Ku NO, Zhou X, Toivola DM, Omary MB. The cytoskeleton of digestive epithelia in health and disease. Am J Physiol 1999; 277:G1108-37.
[163] Chasan B, Geisse NA, Pedatella K, Wooster DG, Teintze M*, et al.* Evidence for direct interaction between actin and the cystic fibrosis transmembrane conductance regulator. Eur Biophys J 2002; 30:617-24.
[164] Lindahl M, Stahlbom B, Tagesson C. Identification of a new potential airway irritation marker, palate lung nasal epithelial clone protein, in human nasal lavage fluid with two-dimensional electrophoresis and matrix-assisted laser desorption/ionization-time of flight. Electrophoresis 2001; 22:1795-800.
[165] Campos MA, Abreu AR, Nlend MC, Cobas MA, Conner GE*, et al.* Purification and characterization of PLUNC from human tracheobronchial secretions. Am J Respir Cell Mol Biol 2004; 30:184-92.
[166] Di YP, Harper R, Zhao Y, Pahlavan N, Finkbeiner W*, et al.* Molecular cloning and characterization of spurt, a human novel gene that is retinoic acid-inducible and encodes a secretory protein specific in upper respiratory tracts. J Biol Chem 2003; 278:1165-73.
[167] Garcia-Caballero A, Rasmussen JE, Gaillard E, Watson MJ, Olsen JC*, et al.* SPLUNC1 regulates airway surface liquid volume by protecting ENaC from proteolytic cleavage. Proc Natl Acad Sci U S A 2009; 106:11412-7.
[168] Gakhar L, Bartlett JA, Penterman J, Mizrachi D, Singh PK*, et al.* PLUNC is a novel airway surfactant protein with anti-biofilm activity. PLoS One 2010; 5:e9098.
[169] Singh PK, Parsek MR, Greenberg EP, Welsh MJ. A component of innate immunity prevents bacterial biofilm development. Nature 2002; 417:552-5.
[170] Wood LG, Gibson PG, Garg ML. Biomarkers of lipid peroxidation, airway inflammation and asthma. Eur Respir J 2003; 21:177-86.
[171] Montuschi P, Kharitonov SA, Ciabattoni G, Corradi M, van Rensen L*, et al.* Exhaled 8-isoprostane as a new non-invasive biomarker of oxidative stress in cystic fibrosis. Thorax 2000; 55:205-9.

[172] Kettle AJ, Chan T, Osberg I, Senthilmohan R, Chapman AL*, et al.* Myeloperoxidase and protein oxidation in the airways of young children with cystic fibrosis. Am J Respir Crit Care Med 2004; 170:1317-23.
[173] Van Der Vliet A, Nguyen MN, Shigenaga MK, Eiserich JP, Marelich GP*, et al.* Myeloperoxidase and protein oxidation in cystic fibrosis. Am J Physiol Lung Cell Mol Physiol 2000; 279:L537-46.
[174] Starosta V, Rietschel E, Paul K, Baumann U, Griese M. Oxidative changes of bronchoalveolar proteins in cystic fibrosis. Chest 2006; 129:431-7.
[175] Linsdell P, Hanrahan JW. Glutathione permeability of CFTR. Am J Physiol 1998; 275:C323-6.
[176] Kogan I, Ramjeesingh M, Li C, Kidd JF, Wang Y*, et al.* CFTR directly mediates nucleotide-regulated glutathione flux. Embo J 2003; 22:1981-9.
[177] Roum JH, Buhl R, McElvaney NG, Borok Z, Crystal RG. Systemic deficiency of glutathione in cystic fibrosis. J Appl Physiol 1993; 75:2419-24.
[178] Wang Y, Phelan SA, Manevich Y, Feinstein SI, Fisher AB. Transgenic mice overexpressing peroxiredoxin 6 show increased resistance to lung injury in hyperoxia. Am J Respir Cell Mol Biol 2006; 34:481-6.
[179] Fujii T, Fujii J, Taniguchi N. Augmented expression of peroxiredoxin VI in rat lung and kidney after birth implies an antioxidative role. Eur J Biochem 2001; 268:218-25.
[180] Trudel S, Kelly M, Fritsch J, Nguyen-Khoa T, Therond P*, et al.* Peroxiredoxin 6 fails to limit phospholipid peroxidation in lung from Cftr-knockout mice subjected to oxidative challenge. PLoS One 2009; 4:e6075.
[181] Arthur JR. The glutathione peroxidases. Cell Mol Life Sci 2000; 57:1825-35.
[182] Strange RC, Spiteri MA, Ramachandran S, Fryer AA. Glutathione-S-transferase family of enzymes. Mutat Res 2001; 482:21-6.
[183] Gomes-Alves P, Couto F, Pesquita C, Coelho AV, Penque D. Rescue of F508del-CFTR by RXR motif inactivation triggers proteome modulation associated with the unfolded protein response. Biochim Biophys Acta 2010; 1804:856-865.
[184] Jefferies H, Coster J, Khalil A, Bot J, McCauley RD*, et al.* Glutathione. ANZ J Surg 2003; 73:517-22.
[185] Wang W, Xu H, Shi Y, Nandedkar S, Zhang H*, et al.* Genetic deletion of apolipoprotein A-I increases airway hyperresponsiveness, inflammation, and collagen deposition in the lung. J Lipid Res 2010; 51:2560-70.
[186] Charro N, Hood BL, Faria D, Pacheco P, Azevedo P*, et al.* Serum proteomics signature of Cystic Fibrosis patients: A complementary 2-DE and LC-MS/MS approach. J Proteomics 2011.
[187] Duval C, Muller M, Kersten S. PPARalpha and dyslipidemia. Biochim Biophys Acta 2007; 1771:961-71.
[188] Reynders V, Loitsch S, Steinhauer C, Wagner T, Steinhilber D*, et al.* Peroxisome proliferator-activated receptor alpha (PPAR alpha) down-regulation in cystic fibrosis lymphocytes. Respir Res 2006; 7:104.
[189] Bunger M, van den Bosch HM, van der Meijde J, Kersten S, Hooiveld GJ*, et al.* Genome-wide analysis of PPARalpha activation in murine small intestine. Physiol Genomics 2007; 30:192-204.
[190] Bolanos JP, Almeida A, Moncada S. Glycolysis: a bioenergetic or a survival pathway? Trends Biochem Sci 2010; 35:145-9.
[191] Bolanos JP, Almeida A. The pentose-phosphate pathway in neuronal survival against nitrosative stress. IUBMB Life 2010; 62:14-8.

Chapter VI

General Discussion and Concluding Remarks

Cystic Fibrosis (CF) remains the most lethal monogenic recessive disease in Caucasians without an effective therapy yet. Classically, it is characterized by a triad of symptoms such as elevated sweat chloride concentrations, pancreatic insufficiency and severe obstruction and bacterial infection of the airways, this latter constituting the most life-compromising condition among patients. The increase of knowledge about CF has exponentially grown since the discovery of its causing gene, the CF Conductance Transmembrane Regulator (CFTR), in 1989 by Lap-Chee Tsui's group [1-3]. This remarkable discovery opened the possibility to uncover the molecular basis of the disease while giving the opportunity to explore other molecular determinants which might contribute to the development and/or progression of CF. However, the quest is far from being ended...

Proteomic methodologies have the potential to uncover hundreds to thousands of proteins in biological samples and comparative analysis allows the identification of differentially expressed specimens across two or more conditions. Application of proteomics to the CF's problematic has delivered important information on the proteostasis involved in the development of the disease, from CFTR interactors to CF-modifier genes, passing by the processes of synthesis, maturation and delivery of CFTR to the plasma membrane. The final aspiration of such studies is to provide validated a protein or patterns of proteins that can be used as specific biomarkers of CF while distinguishing it from other symptoms-sharing diseases, such as COPD and asthma.

The work presented in this thesis comes from a natural path of investigation in our group: we initially started studies on cell lines [4,5], moved forward to a CF animal model [6] and finally investigated human biological samples [7,8]. Here, we have analysed serum, red blood cells and nasal epithelial cells of CF patients in comparison with healthy CF-carriers and non-CF individuals by distinct proteomics methodologies aiming to provide extra knowledge about the proteinaceous environment that characterizes the disease.

When using human samples, reliable results are only obtained when a close collaboration with clinicians and pathologists exist. Criterious and rigorous patients' evaluation and samples' collection and processing under standard and controlled conditions were primordial to give extra confidence on the results presented.

Chapter II focuses in the comparison of serum proteins' abundance of CF patients, CF-carriers and healthy non-CF individuals by complementary gel-based 2-DE-MALDI-TOF/TOF MS and gel-free LC-MS/MS strategies. This approach was due to complementary studies under investigation in the laboratory but also due to serum's high value in terms of biomarkers discovery, as it virtually contains proteins from every tissue or organ of the body, and ease of collection and potential implementation of a clinical test from the results obtainable. Individual serum samples were grouped according with to their genetic background to emphasize proteomics differences while diluting potential individual contributions aiming to praise significant differences between CF and non-CF groups. Serum's immunodepletion was performed in order to reduce proteomics complexity before protein separation and identification by 2-DE-MALDI-TOF/TOF MS and LC-MS/MS. Proteins with altered abundances among the groups in study were further characterized using several bioinformatic tools (PIKE, IPA, ProteiOn, Reactome) to retrieve the maximum biological information and functional significance. From the gel-based approach, 78 differentially expressed spots were identified yielding 28 unique identifications and postulating the existence of PTM or protein isoforms. Such hypothesis was investigated by using FindMod and it was demonstrated that some proteins showed high probability of being modified. Overall, these results indicate that the patterns of modifications might be responsible for different functions and/or activities of proteins with virtual relevance in the disease phenotype. The label-free 1D-RPLC-MS/MS strategy provided 569 different proteins identified by at least 2 peptides, a significantly higher number of identifications than 2-DE, as expected. Pathway analysis allocated these proteins to processes highly significant in CF's development such as the proinflammatory pathway where NFkB plays a central role (members of the apolipoproteins and complement system family, CD180, BAG4, TNF ligand superfamily member 9, α-1-antitrypsin, regulator of G protein signaling 3, hemopexin, vinculin, HSP70, HSPA8, periaxin), tissue remodeling and protease/antiprotease imbalance (vinculin, radixin, tropomyison IV, paladin, myosin VIIa, dedicator to cytokinesis 1, α-1-antitrypsin, inter-α-trypsin inhibitor heavy chains H3 and H4), innate immune dysfunction (members of the complement system family: C5, C9, C1s subcomponent, C3, C3/C5 convertase, C6), oxidative stress and nutritional status (vitamin D-binding protein, plasma retinol binding protein, transferrin, transthyretin, several apolipoproteins) or inflammation (serum amyloid A and A4 proteins, endothelin-A receptor). Importantly, it was possible to identify BiP, BiP-

associated protein and NDKB desregulated only in the CF group (this latter validated by WB), proteins that are already associated with trafficking, degradation and function of CFTR. Also, NDKB is involved in regulation of pancreatic and lung secretions and neutrophil-mediated inflammation.

Altogether, the results obtained here show processes and proteins implicated in the pathogenesis of CF that are corroborated by several studies reported in the literature. Our proteomic "shotgun" approach demonstrated to be efficient in the characterization of CF-associated proteome profile.

Chapter III reports the study of proteins from mRBC. Several studies have delivered lists of proteins resulting from proteomic analysis of RBC though its significance and influence in the context of CF is scarce. Our own group has attempted the study of mRBC from CF patients in comparison with non-CF by 2D-PAGE but improvements in sample preparation and proteins' separation are still required. The undoubtedly demonstration that CFTR is present in RBC was published in 1998 and a more recent study could actually count the number of CFTR molecules in these cells, also showing a reduction by 4-fold of CFTR in RBC of CF patients when compared with healthy individuals. It is however known that not only CFTR is responsible for the phenotypic changes observed in patients. Other proteins might concur to such alterations. Pulmonary hypertension is observed in some CF patients potentially caused by failure in NO synthesis resulting from deficient deformation-induced ATP release of CF's RBC. We used a methanol-facilitated proteins' extraction combined with label-free 2D-LC-MS/MS to identify differences in abundance of membrane proteins from RBC isolated from peripheral blood of CF patients (with mild or severe pulmonary disease) in comparison with healthy CF-carriers and non-CF individuals. Bioinformatic analysis showed proteins involved in cytoskeleton rearrangements (CTSG, FLOT1, FLOT2, STOM, spectrins, ANK1, ACTB, GYPA, GYPC, EPB41, EPB42), antioxidant (CAT, MPO, EPX, PRDX1, PRDX2) and transport activities (SLC4A1, SLC4A2, SLC2A1, SLC2A4, SLC16A1, SLC2A14, HBD, HBG1, HBA1, HBB) as highly significant in processes taking place in CF. Importantly, several proteins from the cytoskeleton, the spectrin-actin network and linker proteins presented a tendency of lower abundance in CF patients when compared with the other groups: FLOT1, FLOT2, GYPA, GYPC, STPA1, STPB, ACTB and ADD2. These differences in abundances of

proteins from the cytoskeleton and spectrins' network might be the basis of imbalanced and impaired organization of the RBC scaffold with consequences in terms of RBC mechanical deformability and ability to efficiently pass through blood vessels' walls to perform their function. We therefore assume that lack of functional and malleable RBC is in the basis of reduced deformation-induced ATP release and consequent increase in pulmonary hypertension following deterioration of pulmonary function.

Chapters IV and V report results of proteome-wide characterization of NEC using a high-throughput strategy that relies on cell fractionation coupled with MudPIT with possible implications on biomarker and/or therapeutic target discovery for respiratory diseases, including CF. Given the physiological similarities between upper and lower airway epithelia and the notion that reactions in the nose influence or reflect complications in lungs [9-12], NEC have been extensively used to assess airway diseases, such as CF. However, the proteomics investigation presented herein is the most complete one so ever achieved regarding the proteome characterization of NEC while comparing such results with others obtained from bronchial epithelial cells. In total, more that 7000 proteins were identified regardless the correspondent subproteome. However, to achieve higher confidence and accuracy in the results, we only considered proteins identified by at least 2 unique triptic peptides identified in three independent runs, resulting in a final list of 1482 proteins representative of NEC: 535 proteins in the soluble subproteome, 1169 proteins in the membrane subproteome and 222 overlapping the two fractions. Further functional characterization was obtained by resorting to panoply of databases and tools. 947 of the 1482 proteins (63.9%) were annotated as membrane or membrane-associated proteins in GO databases with GRAVY, TMD predictions and HPRD confirming such observations. The contribution of each protein to the total protein content in NEC was estimated by emPAI since this index has shown a high correlation with the actual protein amount in complex mixtures with a wide dynamic range. The most abundant proteins were ranked and analysed concerning the biological processes where they participate in showing a prevalence of mitochondrial, ribosomal and cytoskeleton proteins. Similarly, immunological, molecular transport, morphology or processes related to cellular assembly and organization were associated to the lower abundant proteins.

Ingenuity pathway analysis revealed an enrichment of molecular and cellular functions associated with cell death, protein folding and drug metabolism in soluble fraction while in membrane fraction there was a predominance of functions associated with molecular transport, protein trafficking and cell-to-cell signalling and interaction.

To study the representativeness of NEC as surrogate of the lower airways, our data was compared with a proteomics analysis of bronchial epithelium also collected by harvest brushing and highlighted 517 proteins commonly expressed in airways' epithelia whose main functions were associated with maintenance of physical barrier and immunological defence. In our datasets, it was possible to allocated 141 proteins to the development and/or progression of several respiratory diseases such as lung cancer, severe acute respiratory syndrome (SARS) (ACSL1, ACTN1, ANPEP, BPI, CAMP, CD9, CD63, CLU, DEFA1, FLOT2, GAPDH, GLUL, GSTO1, HBG1, HIST1H1C, HP, ITGAM, LCN2, LTA4H, LTF, MNDA, MTX1, PDXK, RAB13, S100A9, S100P, SERPINA1, TALDO1, TCIRG1, TKT, TUBB2C) or pneumonitis (ACO2, ACTB, ALB, ALDH2, ALDOA, ANXA3, CAMP, ENO1, HSPA5, HTT, MYH9, NOS2, P4HB, PDIA3, PRDX1, PRDX6, SELENBP1, STAT3, TKT, TOP2B, TPI1, TUBB). Two proteins, NOS2 and TPI1, have been associated with asthma while MPO, LTF, ENO1 and SELENBP1 have been associated to the progression of CF.

The conjugation of all these observations supports the hypothesis that NEC reflect certain aspects of lower airways and thus might contribute to extend our knowledge on airway epithelium phenotype allowing further comparative analysis of airway-specific proteome expression in pathological conditions such as CF.

The information contained on chapter IV was used to further investigate putative alterations in the NEC proteome as consequence of CF and explore whether these alterations reflect the lung function impairment in this disease and reported in chapter V. NEC from health individuals (non-CF and CF carriers) and CF patients were obtained and analysed as described in chapter IV. The severity of CF's lung disease assessed by clinical and respiratory function evaluation was also contemplated. Data normalization was performed according to the Total Signal (TS) method, a strategy described by Carvalho *et al.* [13] that takes into account all the peptides identified in an individual run and uses it to normalize the abundance of each protein across runs. After normalization and comparison of the normalized TPC, 512 proteins presented altered

abundances in CF's NEC in comparison with the above mentioned groups. In accordance with previously reported results in serum and mRBC, tissue remodeling, cytoskeleton rearrangements, antioxidant status and inflammation are again among the most represented functions in which differentially expressed proteins participate in. Also, imbalance mitochondrial function, energy production, glucose's metabolism, antigen presentation, protein synthesis and folding and degradation of extracelular matrix were also indicated. Several CFTR-interacting proteins and CF-modifier genes described here have putative roles in the development and/or progression of CF lung disease. Integration of these proteins in pathways and networks with biological significance corroborated our findings at the protein level, retrieving known impaired functions already associated with CF.

We believe that our findings constitute a solid base for more focused and dedicated studies on relevant proteins highlighted here that ultimately can constitute potential therapeutic targets and/or auxiliary diagnostic biomarkers for CF while effectively distinguish it from other respiratory diseases.

Concluding Remarks and Future Directions

The work in the CF's field is far from being ended. Proteomics is a very powerful scientific discipline whose field of application is virtually unlimited, delivering massive amounts of data about the question under study. Application of proteomics technologies to the CF's problematic has already proved to be very efficient and delivering important features that assist in the understanding of CF's pathology.

The profiling and identification of differentially expressed proteins in the several biological samples investigated in the context of this thesis have provided new hints and information on processes imbalanced in CF. Comparing the results obtained and presented along the chapters, it is possible to observe that tissue remodelling, antioxidant status and inflammation are desregulated processes transversal to all different CF samples (serum, mRBC, NEC) here, among others. Confirming the involvement of the proteins contributing to their promotion should be further investigated. Once again, proteomics can be one of the tools to aid their solution. Validation of some of the specimens identified in these works in individual samples from a larger cohort of patients and including samples from other CF symptoms-sharing diseases could be very beneficial in identifying specific targets from therapeutic intervention. This validation can be performed, for example, by Western Blot, or by taking advantage of some powerful MS capabilities, such as single or multiple reaction monitoring (SRM or MRM, respectively) [14]. In fact, some applications for S/MRM are already published (for review, see [15]) aiming to validate biomarkers previously identified in the discovery phase. Compared with WB or ELISA, it is cheaper to set up a MRM assays in an automated fashion and to target multiple protein targets simultaneously or monitor multiple product ions for each peptide precursor from single protein. The only requirements for a successful S/MRM assays is the selection of the appropriate transitions, *i.e.*, those that present high relative intensity and sensitivity m/z peaks for the precursor and product ion(s) and that are unique of a certain protein, with unique elution time in HLPC and precursor/product m/z values specific to the protein in the assay.

Being associated with the National Institute of Health, some assays used for routine clinical diagnosis could be applied to the validation of our results under certificate, reproducible and high quality methodology. Some of the results obtained here were already validated by immunoturbidimety; others are under evaluation by other regular clinical tests, allowing the processing of hundreds of samples in a controlled manner. To take advantage of these clinical routine methodologies to validate our proteomics data in all patients' samples individually seems a very promising approach.

In one of the works presented we postulated that PTM might be in the basis for different severities of lung disease observed among patients and relevant for the disease progression. Confirming these results by the enrichment of specific modified proteins or employing MS methods to sequence and determine the type of PTM could help in answering to this question.

While new perspectives on CF were brought up with this investigation, other questions remained to be answered. Ultimately, our efforts go in the direction of providing new auxiliary targets for CF's diagnostic/prognostic or targets for the development of therapeutic agents for improvements in life's quality and increase in life's expectancy. To achieve that, specific molecules must be uncovered. This study will benefit with the inclusion of other pulmonary diseases that share some of the CF's phenotype for comparison purposes, such as asthma and Chronic Obstructive Pulmonary Disease (COPD). At the same time, investigators in the proteomics field faced each day with an impressive amount of data aiming to retrieve the maximum of biological information. Databases and bioinformatic tools are already providing a good help in that field but more developments and improvements are necessary to accompany the rapid expansion and progress of the technical components that Proteomics makes use of.

REFERENCES

[1] Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G*, et al.* Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 1989; 245:1059-65.

[2] Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R*, et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989; 245:1066-73.
[3] Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK*, et al.* Identification of the cystic fibrosis gene: genetic analysis. Science 1989; 245:1073-80.
[4] Gomes-Alves P, Couto F, Pesquita C, Coelho AV, Penque D. Rescue of F508del-CFTR by RXR motif inactivation triggers proteome modulation associated with the unfolded protein response. Biochim Biophys Acta 2010; 1804:856-865.
[5] Gomes-Alves P, Neves S, Coelho AV, Penque D. Low temperature restoring effect on F508del-CFTR misprocessing: A proteomic approach. J Proteomics 2009; 73:218-30.
[6] Carvalho-Oliveira IM, Charro N, Aarbiou J, Buijs-Offerman RM, Wilke M*, et al.* Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009; 8:3606-16.
[7] Gomes-Alves P, Imrie M, Gray RD, Nogueira P, Ciordia S*, et al.* SELDI-TOF biomarker signatures for cystic fibrosis, asthma and chronic obstructive pulmonary disease. Clin Biochem 2010; 43:168-77.
[8] Charro N, Hood BL, Faria D, Pacheco P, Azevedo P*, et al.* Serum proteomics signature of Cystic Fibrosis patients: A complementary 2-DE and LC-MS/MS approach. J Proteomics 2011.
[9] Passalacqua G, Ciprandi G, Canonica GW. The nose-lung interaction in allergic rhinitis and asthma: united airways disease. Curr Opin Allergy Clin Immunol 2001; 1:7-13.
[10] Serrano C, Valero A, Picado C. [Rhinitis and asthma: one airway, one disease]. Arch Bronconeumol 2005; 41:569-78.
[11] Rimmer J, Ruhno JW. 6: Rhinitis and asthma: united airway disease. Med J Aust 2006; 185:565-71.
[12] Braunstahl GJ. The united airways concept: from bench to bedside. Monaldi Arch Chest Dis 2007; 67:95-101.
[13] Carvalho PC, Hewel J, Barbosa VC, Yates JR, 3rd. Identifying differences in protein expression levels by spectral counting and feature selection. Genet Mol Res 2008; 7:342-56.
[14] Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol 2008; 4:222.
[15] Cham Mead JA, Bianco L, Bessant C. Free computational resources for designing selected reaction monitoring transitions. Proteomics 2010; 10:1106-26.